WEST Search History

Hide Items Restore Clear Cancel

DATE: Monday, January 12, 2004

Hide?	**Set Name**	**Query**	**Hit Count**
		DB=PGPB,USPT,EPAB,JPAB,DWPI; PLUR=YES; OP=OR	
☐	L7	L6 not (l4 or l5)	3
☐	L6	l3 and l2	3
☐	L5	zaremba-sam$.in.	2
☐	L4	barzaga-elene$.in.	1
☐	L3	schlom-jeffrey$.in.	32
☐	L2	(cea or (carcinoembroyic adj antigen))adj5 vaccine$	29
☐	L1	cea and vaccine$	1139

END OF SEARCH HISTORY

```
                GenCore version 5.1.6
         Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         January 12, 2004, 14:25:20 ; Search time 10.25 Seconds
                                           (without alignments)
                                           84.441 Million cell updates/sec

Title:          US-09-529-121A-2
Perfect score:  45
Sequence:       1 YLSGADLNL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters:      789

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_76:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                     %
Result               Query
  No.   Score  Match Length DB  ID                  Description
-----------------------------------------------------------------------
    1      16   35.6      7  2  PT0542              T-cell receptor be
    2      16   35.6      7  2  PC2370              probable H+-transp
    3      16   35.6      8  2  JS0318              leucokinin VIII -
    4      16   35.6      9  1  AKLQIM              locustamyoinhibiti
    5      16   35.6      9  2  A57444              neuropeptide Grb-A
    6      16   35.6      9  2  B57444              neuropeptide Grb-A
    7      16   35.6      9  4  I57650              hemoglobin alpha c
    8      15   33.3      5  2  PT0540              T-cell receptor be
    9      15   33.3      6  2  PT0726              T-cell receptor be
   10      15   33.3      7  2  PT0526              T-cell receptor be
   11      15   33.3      7  2  PT0676              T-cell receptor be
   12      15   33.3      8  2  T13818              cytochrome oxidase
   13      15   33.3      9  2  F41978              calliFMRFamide 6 -
   14      15   33.3      9  2  PT0288              Ig heavy chain CRD
   15      15   33.3      9  2  G41946              T-cell receptor ga
   16      14   31.1      4  2  A26209              protein-glutamine
   17      14   31.1      5  2  PT0679              T-cell receptor be
   18      14   31.1      6  2  PT0605              T-cell receptor be
   19      14   31.1      6  2  PT0593              T-cell receptor be
   20      14   31.1      7  2  S20446              elastase - Pseudom
   21      14   31.1      7  2  PT0654              T-cell receptor be
   22      14   31.1      7  2  PT0722              T-cell receptor be
   23      14   31.1      8  2  A21440              variant surface gl
   24      14   31.1      8  2  A41117              acetylcholinestera
   25      14   31.1      8  2  PN0043              phosphatidylethano
   26      14   31.1      8  2  PT0557              T-cell receptor be
   27      14   31.1      9  2  A61364              isotocin - common
   28      14   31.1      9  2  C57444              neuropeptide Grb-A
   29      13   28.9      4  2  S43959              Ig mu chain V regi
   30      13   28.9      7  2  I50210              gene c-rel protein
   31      13   28.9      8  2  PT0547              T-cell receptor be
   32      13   28.9      9  2  A44873              caldesmon - rabbit
   33      13   28.9      9  2  QDRB                delta sleep-induci
   34      13   28.9      9  2  C41170              photosystem II pro
   35      13   28.9      9  2  PH0935              T-cell receptor be
   36      13   28.9      9  2  PH0918              T-cell receptor be
   37      12   26.7      6  2  I51434              H4 histone - Afric
   38      12   26.7      7  2  S16364              opacity protein P.
   39      12   26.7      7  2  B35890              RNA-directed DNA p
   40      12   26.7      7  2  PN0649              pullulanase (EC 3.
   41      12   26.7      7  2  S29735              polyphosphate-gluc
   42      12   26.7      8  2  PQ0012              cholecystokinin -
   43      12   26.7      8  2  A43001              cholecystokinin -
   44      12   26.7      8  2  PL0184              capsid protein VP-
   45      12   26.7      8  2  S65647              2-hydroxyglutaryl-

                                  ALIGNMENTS

RESULT 1
PT0542
T-cell receptor beta chain V-D-J region (126-1BA) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0542
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0542
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: day 18 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             35.6%;  Score 16;  DB 2;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              || |
Db          2 SGGD 5

RESULT 2
PC2370
probable H+-transporting two-sector ATPase (EC 3.6.3.14) alpha chain [similarity] - Baci
N;Alternate names: unidentified 78K protein
C;Species: Bacillus cereus
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 03-Jun-2002
C;Accession: PC2370
R;Matsuno, K.; Miyamoto, T.; Yamaguchi, K.; Sayed, M.A.; Kajiwara, T.; Hatano, S.
Biosci. Biotechnol. Biochem. 59, 231-235, 1995
A;Title: Identification of DNA-binding proteins changed after induction of sporulation i
A;Reference number: PC2369; MUID:95218265; PMID:7766022
A;Accession: PC2370
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MAS>
C;Keywords: ATP biosynthesis; hydrolase

  Query Match             35.6%;  Score 16;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DLN 8
              |||
Db          2 DLN 4
```

```
RESULT 3
JS0318
leucokinin VIII - Madeira cockroach
C;Species: Leucophaea maderae (Madeira cockroach)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Jun-2000
C;Accession: JS0318
R;Holman, G.M.; Cook, B.J.; Nachman, R.J.
Comp. Biochem. Physiol. C 88, 31-34, 1987
A;Title: Isolation, primary structure and synthesis of leucokinins VII and VIII: the fir
A;Reference number: JS0317
A;Accession: JS0318
A;Molecule type: protein
A;Residues: 1-8 <HOL>
C;Comment: Leucokinins, a family of cephalomyotropic peptides, stimulate contractile act
C;Keywords: amidated carboxyl end; cephalomyotropic peptide
F;8/Modified site: amidated carboxyl end (Gly) #status experimental

  Query Match             35.6%;  Score 16;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GAD 6
              |||
Db          1 GAD 3

RESULT 4
AKLQIM
locustamyoinhibiting peptide - migratory locust
C;Species: Locusta migratoria (migratory locust)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 20-Mar-1998
C;Accession: A60065
R;Schoofs, L.; Holman, G.M.; Hayes, T.K.; Nachman, R.J.; De Loof, A.
Regul. Pept. 36, 111-119, 1991
A;Title: Isolation, identification and synthesis of locustamyoinhibiting peptide (LOM-MI
A;Reference number: A60065; MUID:92179466; PMID:1796179
A;Accession: A60065
A;Molecule type: protein
A;Residues: 1-9 <SCH>
C;Comment: This peptide hormone suppresses spontaneous contractions of the hindgut and c
C;Superfamily: locustamyoinhibiting peptide
C;Keywords: amidated carboxyl end; hormone
F;9/Modified site: amidated carboxyl end (Trp) #status experimental

  Query Match             35.6%;  Score 16;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DLN 8
              |||
Db          4 DLN 6

RESULT 5
A57444
neuropeptide Grb-AST B1 - two-spotted cricket
C;Species: Gryllus bimaculatus (two-spotted cricket)
C;Date: 26-Jan-1996 #sequence_revision 26-Jan-1996 #text_change 26-Jan-1996
C;Accession: A57444
R;Lorenz, M.W.; Kellner, R.; Hoffmann, K.H.
J. Biol. Chem. 270, 21103-21108, 1995
A;Title: A family of neuropeptides that inhibit juvenile hormone biosynthesis in the cri
A;Reference number: A57444; MUID:95403341; PMID:7673141
A;Accession: A57444
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <LOR>

  Query Match             35.6%;  Score 16;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DLN 8
              |||
Db          4 DLN 6

RESULT 6
B57444
neuropeptide Grb-AST B2 - two-spotted cricket
C;Species: Gryllus bimaculatus (two-spotted cricket)
C;Date: 26-Jan-1996 #sequence_revision 26-Jan-1996 #text_change 26-Jan-1996
C;Accession: B57444
R;Lorenz, M.W.; Kellner, R.; Hoffmann, K.H.
J. Biol. Chem. 270, 21103-21108, 1995
A;Title: A family of neuropeptides that inhibit juvenile hormone biosynthesis in the cric
A;Reference number: A57444; MUID:95403341; PMID:7673141
A;Accession: B57444
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <LOR>

  Query Match             35.6%;  Score 16;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DLN 8
              |||
Db          4 DLN 6

RESULT 7
I57650
hemoglobin alpha chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 31-Jul-1997 #text_change 20-Apr-2000
C;Accession: I57650
R;Whitelaw, E.; Hogben, P.; Hanscombe, O.; Proudfoot, N.J.
Mol. Cell. Biol. 9, 241-251, 1989
A;Title: Transcriptional promiscuity of the human alpha-globin gene.
A;Reference number: I57650; MUID:89181576; PMID:2538719
A;Accession: I57650
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-9 <WHI>
A;Cross-references: GB:M23454; NID:g340922; PIDN:AAA52629.1; PID:g553329
A;Note: engineered sequence; this sequence was not determined in this report
C;Genetics:
A;Gene: GDB:HBA1
A;Cross-references: GDB:119293
A;Map position: 16p13.3-16p13.3

  Query Match             35.6%;  Score 16;  DB 4;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 2.8e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 LSGAD 6
              || ||
Db          3 LSPAD 7

RESULT 8
PT0540
T-cell receptor beta chain V-D-J region (126-1L) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0540
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0540
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: day 18 fetal thymus, strain BALB/c
```

```
C;Keywords: T-cell receptor

  Query Match             33.3%;  Score 15;  DB 2;  Length 5;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              || |
Db          2 SGED 5


RESULT 9
PT0726
T-cell receptor beta chain V-D-J region (161-2D) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0726
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0726
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-6 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             33.3%;  Score 15;  DB 2;  Length 6;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              || |
Db          2 SGED 5


RESULT 10
PT0526
T-cell receptor beta chain V-D-J region (100-4E) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0526
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0526
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: adult thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             33.3%;  Score 15;  DB 2;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              || |
Db          2 SGED 5

RESULT 11
PT0676
T-cell receptor beta chain V-D-J region (140-1AL) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0676
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
```

```
A;Accession: PT0676
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-7 <FEE>
A;Experimental source: day 18 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             33.3%;  Score 15;  DB 2;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              || |
Db          2 SGED 5


RESULT 12
T13818
cytochrome oxidase subunit I - Atlantic hagfish mitochondrion (fragment)
C;Species: mitochondrion Myxine glutinosa (Atlantic hagfish)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T13818
R;Delarbre, C.; Barriel, V.; Tillier, S.; Janvier, P.; Gachelin, G.
Mol. Biol. Evol. 14, 807-813, 1997
A;Title: The main features of the craniate mitochondrial DNA between the ND1 and the COI
A;Reference number: Z17775; MUID:97398704; PMID:9254918
A;Accession: T13818
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-8 <DEL>
A;Cross-references: EMBL:Y09527; NID:g2340019; PIDN:CAA70718.1; PID:g2340022
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion

  Query Match             33.3%;  Score 15;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLS 3
              |||
Db          2 YLS 4


RESULT 13
F41978
calliFMRFamide 6 - bluebottle fly (Calliphora vomitoria)
C;Species: Calliphora vomitoria
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
C;Accession: F41978
R;Duve, H.; Johnsen, A.H.; Sewell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, J.F.; Thorpe,
Proc. Natl. Acad. Sci. U.S.A. 89, 2326-2330, 1992
A;Title: Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH-2 neuropeptides (desig
A;Reference number: A41978; MUID:92196111; PMID:1549595
A;Accession: F41978
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <DUV>
C;Keywords: amidated carboxyl end; neuropeptide
F;9/Modified site: amidated carboxyl end (Phe) #status experimental

  Query Match             33.3%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              || |
Db          2 SGQD 5


RESULT 14
PT0288
```

```
Ig heavy chain CRD3 region (clone 4-106) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0288
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0288
A;Molecule type: DNA
A;Residues: 1-9 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

  Query Match             33.3%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSG 4
              | ||
Db          5 YSSG 8


RESULT 15
G41946
T-cell receptor gamma chain (2t.23) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C;Accession: G41946
R;Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A;Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
A;Reference number: A41946; MUID:92049316; PMID:1658619
A;Accession: G41946
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-9 <WHE>
C;Keywords: T-cell receptor

  Query Match             33.3%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSG 4
              | ||
Db          5 YSSG 8


Search completed: January 12, 2004, 14:31:52
Job time : 12.25 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         January 12, 2004, 14:21:34 ; Search time 6.25 Seconds
                                             (without alignments)
                                             67.718 Million cell updates/sec

Title:          US-09-529-121A-2
Perfect score:  45
Sequence:       1 YLSGADLNL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters:      251

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_41:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                   %
Result           Query
  No.   Score    Match Length DB  ID                      Description
-------------------------------------------------------------------------
    1     17     37.8      8  1  CPD1_ENTFA              P13269 enterococcu
    2     16     35.6      8  1  LCK8_LEUMA              P19990 leucophaea
    3     16     35.6      9  1  LMIP_LOCMI              P31799 locusta mig
    4     16     35.6      9  1  PTSP_BOMMO              P82003 bombyx mori
    5     15     33.3      9  1  FAR6_CALVO              P41861 calliphora
    6     14     31.1      8  1  ALL5_CYDPO              P82156 cydia pomon
    7     14     31.1      8  1  FAR2_MACRS              P83275 macrobrachi
    8     14     31.1      9  1  ISOT_CYPCA              P42993 cyprinus ca
    9     13     28.9      8  1  PLP_BRANA               P81707 brassica na
   10     13     28.9      9  1  DSIP_RABIT              P01158 oryctolagus
   11     13     28.9      9  1  FAR2_PANRE              P41873 panagrellus
   12     13     28.9      9  1  OXYA_SQUAC              P42999 squalus aca
   13     13     28.9      9  1  OXYT_RAJCL              P42994 raja clavat
   14     12     26.7      7  1  PPH2_LYCES              P83379 lycopersico
   15     12     26.7      8  1  CCKN_MACEU              P30369 macropus eu
   16     12     26.7      9  1  FARD_CALVO              P41868 calliphora
   17     12     26.7      9  1  OXYT_RABIT              P32878 oryctolagus
   18     12     26.7      9  1  PGLR_DIAAB              P81179 diaprepes a
   19     11     24.4      4  1  FAR3_HIRME              P42562 hirudo medi
   20     11     24.4      5  1  PRCT_PERAM              P01373 periplaneta
   21     11     24.4      6  1  CIP2_MYTED              P13737 mytilus edu
   22     11     24.4      7  1  FAR2_ASCSU              P31890 ascaris suu
   23     11     24.4      7  1  GFRP_MOUSE              P99025 mus musculu
   24     11     24.4      7  1  LANC_CARUI              P36960 carnobacter
   25     11     24.4      7  1  UF03_MOUSE              P38641 mus musculu
   26     11     24.4      8  1  CAD1_ENTFA              P13268 enterococcu
   27     11     24.4      8  1  FAR3_HOMAM              P41486 homarus ame
   28     11     24.4      8  1  FAR8_CALVO              P41863 calliphora
   29     11     24.4      9  1  D1_NEPNO                P24816 nephrops no
   30     11     24.4      9  1  FAR5_CALVO              P41860 calliphora
   31     11     24.4      9  1  FAR7_CALVO              P41862 calliphora
   32     11     24.4      9  1  MOSF_CLYJA              P19853 clypeaster
   33     10     22.2      4  1  ACH1_ACHFU              P35904 achatina fu
   34     10     22.2      5  1  UXA4_CHLTR              P38005 chlamydia t
   35     10     22.2      6  1  TMOF_SARBU              P41495 sarcophaga
   36     10     22.2      6  1  TRPI_PSEPU              P36414 pseudomonas
   37     10     22.2      7  1  ALL2_CARMA              P81805 carcinus ma
   38     10     22.2      7  1  ALL3_CARMA              P81806 carcinus ma
   39     10     22.2      7  1  ALL4_CARMA              P81807 carcinus ma
   40     10     22.2      7  1  ALL5_CARMA              P81808 carcinus ma
   41     10     22.2      7  1  ALL7_CYDPO              P82158 cydia pomon
   42     10     22.2      7  1  FAR1_MACRS              P83274 macrobrachi
   43     10     22.2      7  1  FAR2_PROCL              P38498 procambarus
   44     10     22.2      7  1  UN06_PINPS              P81675 pinus pinas
   45     10     22.2      8  1  ACT_CARMA               P80709 carcinus ma

                              ALIGNMENTS

RESULT 1
CPD1_ENTFA
ID   CPD1_ENTFA     STANDARD;      PRT;     8 AA.
AC   P13269;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Sex pheromone cPD1.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=85040388; PubMed=6436978;
RA   Suzuki A., Mori M., Sagakami Y., Isogai A., Fujino M., Kitada C.,
RA   Craig R.A., Clewell D.B.;
RT   "Isolation and structure of bacterial sex pheromone, cPD1.";
RL   Science 226:849-850(1984).
CC   -!- FUNCTION: CPD1 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC       BACTERIOCIN PLASMID PPD1.
KW   Pheromone.
SQ   SEQUENCE   8 AA;  913 MW;  8665B729C682C729 CRC64;

  Query Match             37.8%;  Score 17;  DB 1;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSG 4
              :|||
Db          5 FLSG 8


RESULT 2
LCK8_LEUMA
ID   LCK8_LEUMA     STANDARD;      PRT;     8 AA.
AC   P19990;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Leucokinin VIII (L-VIII).
OS   Leucophaea maderae (Madeira cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
OC   Blaberidae; Leucophaea.
OX   NCBI_TaxID=6988;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Head;
RA   Holman G.M., Cook B.J., Nachman R.J.;
RT   "Isolation, primary structure and synthesis of leucokinins VII and
RT   VIII: the final members of this new family of cephalomyotropic
RT   peptides isolated from head extracts of Leucophaea maderae.";
RL   Comp. Biochem. Physiol. 88C:31-34(1987).
CC   -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC       ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC   -!- SIMILARITY: TO THE OTHER LEUCOKININS.
```

```
DR   PIR; JS0318; JS0318.
KW   Neuropeptide; Amidation.
FT   MOD_RES      8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  902 MW;  736365AB59CAADD8 CRC64;

  Query Match             35.6%;  Score 16;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GAD 6
              |||
Db          1 GAD 3


RESULT 3
LMIP_LOCMI
ID   LMIP_LOCMI     STANDARD;      PRT;     9 AA.
AC   P31799;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-OCT-1993 (Rel. 27, Last annotation update)
DE   Locustamyoinhibiting peptide (LOM-MIP).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC   Acridoidea; Acrididae; Oedipodinae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92179466; PubMed=1796179;
RA   Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;
RT   "Isolation, identification and synthesis of locustamyoinhibiting
RT   peptide (LOM-MIP), a novel biologically active neuropeptide from
RT   Locusta migratoria.";
RL   Regul. Pept. 36:111-119(1991).
CC   -!- FUNCTION: SUPPRESSES SPONTANEOUS CONTRACTIONS OF THE HINDGUT AND
CC       OVIDUCT.
CC   -!- TISSUE SPECIFICITY: NEURONS LOCATED IN TWO VENTRAL CELL CLUSTERS
CC       IN THE SUBOESOPHAGEAL GANGLION.
DR   PIR; A60065; AKLQIM.
KW   Amidation; Neuropeptide.
FT   MOD_RES      9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1060 MW;  387D7DD4472AB6C3 CRC64;

  Query Match             35.6%;  Score 16;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DLN 8
              |||
Db          4 DLN 6


RESULT 4
PTSP_BOMMO
ID   PTSP_BOMMO     STANDARD;      PRT;     9 AA.
AC   P82003;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Prothoracicostatic peptide (Bom-PTSP).
OS   Bombyx mori (Silk moth).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC   Bombycidae; Bombyx.
OX   NCBI_TaxID=7091;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=C145 X N140; TISSUE=Brain;
RX   MEDLINE=20002634; PubMed=10531308;
RA   Hua Y.-J., Tanaka Y., Nakamura K., Sakakibara M., Nagata S.,
RA   Kataoka H.;
RT   "Identification of a prothoracicostatic peptide in the larval brain of
```

```
RT   the silkworm, Bombyx mori.";
RL   J. Biol. Chem. 274:31169-31173(1999).
RN   [2]
RP   ERRATUM.
RA   Hua Y.-J., Tanaka Y., Nakamura K., Sakakibara M., Nagata S.,
RA   Kataoka H.;
RL   J. Biol. Chem. 275:9892-9892(2000).
CC   -!- FUNCTION: Inhibits ecdysteroid biosynthesis in the prothoracic
CC       gland.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DEVELOPMENTAL STAGE: EARLY FIFTH INSTAR.
KW   Hormone; Amidation.
FT   MOD_RES      9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1090 MW;  3878C5B4472AB6C3 CRC64;

  Query Match             35.6%;  Score 16;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DLN 8
              |||
Db          4 DLN 6


RESULT 5
FAR6_CALVO
ID   FAR6_CALVO     STANDARD;      PRT;     9 AA.
AC   P41861;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   CalliFMRFamide 6.
OS   Calliphora vomitoria (Blue blowfly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC   Calliphoridae; Calliphora.
OX   NCBI_TaxID=27454;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Thoracic ganglion;
RX   MEDLINE=92196111; PubMed=1549595;
RA   Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA   Rehfeld J.F., Thorpe A.;
RT   "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT   neuropeptides (designated calliFMRFamides) from the blowfly
RT   Calliphora vomitoria.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
DR   PIR; F41978; F41978.
KW   Neuropeptide; Amidation.
FT   MOD_RES      9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1058 MW;  96D10699CAB6D865 CRC64;

  Query Match             33.3%;  Score 15;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              || |
Db          2 SGQD 5


RESULT 6
ALL5_CYDPO
ID   ALL5_CYDPO     STANDARD;      PRT;     8 AA.
AC   P82156;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Cydiastatin 5.
OS   Cydia pomonella (Codling moth).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
```

```
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Tortricoidea; Tortricidae; Olethreutinae; Cydia.
OX   NCBI_TaxID=82600;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Larva;
RX   MEDLINE=98054539; PubMed=9392829;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA   Davey M., East P.D., Thorpe A.;
RT   "Lepidopteran peptides of the allatostatin superfamily.";
RL   Peptides 18:1301-1309(1997).
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  898 MW;  922879CABB58640D CRC64;

  Query Match             31.1%;  Score 14;  DB 1;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 GADLNL 9
              | |  |
Db          3 GYDFGL 8


RESULT 7
FAR2_MACRS
ID   FAR2_MACRS     STANDARD;      PRT;     8 AA.
AC   P83275;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   FMRFamide-like neuropeptide FLP2 (ADKNFLRF-amide).
OS   Macrobrachium rosenbergii (Giant fresh water prawn).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea;
OC   Palaemonoidea; Palaemonidae; Macrobrachium.
OX   NCBI_TaxID=79674;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=Eyestalk;
RA   Sithigorngul P., Saraithongkum W., Jaideechoey S., Longyant S.,
RA   Sithigorngul W.;
RT   "Novel FMRFamide-like neuropeptides from the eyestalk of the giant
RT   freshwater prawn Macrobrachium rosenbergii.";
RL   Comp. Biochem. Physiol. 120B:587-595(1998).
CC   -!- MASS SPECTROMETRY: MW=1009.4; METHOD=MALDI.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
DR   GO; GO:0007218; P:neuropeptide signaling pathway; TAS.
KW   Neuropeptide; Amidation.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  1010 MW;  9CD40729C4433AAD CRC64;

  Query Match             31.1%;  Score 14;  DB 1;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 ADLN 8
              || |
Db          1 ADKN 4


RESULT 8
ISOT_CYPCA
ID   ISOT_CYPCA     STANDARD;      PRT;     9 AA.
AC   P42993;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Isotocin.
OS   Cyprinus carpio (Common carp).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Cyprinus.
OX   NCBI_TaxID=7962;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Pituitary;
RA   Acher R., Chauvet J., Chauvet M.-T., Crepy D.;
RT   "Characterization of neurohypophyseal hormones from a fresh water bony
RT   fish, the carp (Cyprinus carpio). Comparison with hormones from sea
RT   water bony fishs.";
RL   Comp. Biochem. Physiol. 14:245-254(1965).
CC   -!- FUNCTION: ANTIDIURETIC HORMONE.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   PIR; A61364; A61364.
DR   InterPro; IPR000981; Neurhyp_horm.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Amidation.
FT   DISULFID      1      6
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  969 MW;  17FF476EB455B04B CRC64;

  Query Match             31.1%;  Score 14;  DB 1;  Length 9;
  Best Local Similarity   28.6%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 YLSGADL 7
              |:|   :
Db          2 YISNCPI 8


RESULT 9
PLP_BRANA
ID   PLP_BRANA      STANDARD;      PRT;     8 AA.
AC   P81707;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Plastidial lipid-associated protein (Fragment).
OS   Brassica napus (Rape).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Brassica.
OX   NCBI_TaxID=3708;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=CV. TOPAZ; TISSUE=Tapetum;
RX   MEDLINE=99349136; PubMed=10420651;
RA   Hernandez-Pinzon I., Ross J.H.E., Barnes K.A., Damant A.P.,
RA   Murphy D.J.;
RT   "Composition and role of tapetal lipid bodies in the biogenesis of the
RT   pollen coat of Brassica napus.";
RL   Planta 208:588-598(1999).
CC   -!- FUNCTION: MAY PLAY A STRUCTURAL ROLE IN THE ELAIOPLAST, A TAPETUM-
CC       SPECIFIC PLASTIDIAL LIPID ORGANELLE.
CC   -!- TISSUE SPECIFICITY: TAPETUM OF ANTHERS.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  989 MW;  9D7B1AA452CAA042 CRC64;

  Query Match             28.9%;  Score 13;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DLN 8
              |:|
Db          3 DVN 5


RESULT 10
DSIP_RABIT
ID   DSIP_RABIT     STANDARD;      PRT;     9 AA.
AC   P01158;
DT   21-JUL-1986 (Rel. 01, Created)
```

```
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Delta sleep-inducing peptide (DSIP).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=77185324; PubMed=862769;
RA   Monnier M., Dudler L., Gachter R., Maier P.F., Tobler H.J.,
RA   Schoenenberger G.A.;
RT   "The delta sleep inducing peptide (DSIP). Comparative properties of
RT   the original and synthetic nonapeptide.";
RL   Experientia 33:548-552(1977).
RN   [2]
RP   SEQUENCE, AND SYNTHESIS.
RX   MEDLINE=79054421; PubMed=568769;
RA   Schoenenberger G.A., Maier P.F., Tobler H.J., Wilson K., Monnier M.;
RT   "The delta EEG (sleep)-inducing peptide (DSIP). XI. Amino-acid
RT   analysis, sequence, synthesis and activity of the nonapeptide.";
RL   Pflugers Arch. 376:119-129(1978).
RN   [3]
RP   REVIEW.
RX   MEDLINE=87175129; PubMed=3550726;
RA   Graf M.V., Kastin A.J.;
RT   "Delta-sleep-inducing peptide (DSIP): an update.";
RL   Peptides 7:1165-1187(1986).
CC   -!- FUNCTION: WHEN INFUSED INTO THE MESODIENCEPHALIC VENTRICLE OF
CC       RECIPIENT RABBITS INDUCES SPINDLE AND DELTA EEG ACTIVITY AND
CC       REDUCED MOTOR ACTIVITIES.
CC   -!- MISCELLANEOUS: THIS PEPTIDE WAS OBTAINED FROM DIALYSATES OF
CC       OCCIPITAL VENOUS SINUS BLOOD FROM RABBITS KEPT ASLEEP BY ELECTRIC
CC       STIMULATION OF THE THALAMUS.
CC   -!- DATABASE: NAME=Protein Spotlight;
CC       NOTE=Issue 8 of March 2001;
CC       WWW="http://www.expasy.org/spotlight/articles/sptlt008.html".
DR   PIR; A01422; QDRB.
SQ   SEQUENCE   9 AA;  849 MW;  DDD365BDDAA8787D CRC64;

  Query Match             28.9%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              :| |
Db          2 AGGD 5


RESULT 11
FAR2_PANRE
ID   FAR2_PANRE     STANDARD;      PRT;     9 AA.
AC   P41873;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFamide-like neuropeptide PF2 (SADPNFLRF-amide).
OS   Panagrellus redivivus.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
OC   Panagrolaimoidea; Panagrolaimidae; Panagrellus.
OX   NCBI_TaxID=6233;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=93027659; PubMed=1408999;
RA   Geary T.G., Price D.A., Bowman J.W., Winterrowd C.A., Mackenzie C.D.,
RA   Garrison R.D., Williams J.F., Friedman A.R.;
RT   "Two FMRFamide-like peptides from the free-living nematode
RT   Panagrellus redivivus.";
RL   Peptides 13:209-214(1992).
CC   -!- FUNCTION: MYOACTIVE.
CC   -!- TISSUE SPECIFICITY: NERVE CORDS AND PAIRED GROUPS OF CELLS LOCATED
CC       CAUDALLY TO THE BASE OF THE PHARYNX.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
```

```
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1066 MW;  DA0B0729C4576AAD CRC64;

  Query Match             28.9%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 ADLN 8
              || |
Db          2 ADPN 5


RESULT 12
OXYA_SQUAC
ID   OXYA_SQUAC     STANDARD;      PRT;     9 AA.
AC   P42999;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Aspartocin (Aspargtocin).
OS   Squalus acanthias (Spiny dogfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Squalea; Squaloidei; Squalidae; Squalus.
OX   NCBI_TaxID=7797;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=73031727; PubMed=5083097;
RA   Acher R., Chauvet J., Chauvet M.-T.;
RT   "Phylogeny of the neurohypophysial hormones. Two new active peptides
RT   isolated from a cartilaginous fish, Squalus acanthias.";
RL   Eur. J. Biochem. 29:12-19(1972).
RN   [2]
RP   SEQUENCE.
RX   MEDLINE=72128038; PubMed=4622083;
RA   Acher R., Chauvet J., Chauvet M.-T., Fontaine M.;
RT   "Identification of 2 new neurohypophyseal hormones, valitocin (Val8-
RT   oxytocin) and aspartocin (Asn4-oxytocin) in a selachian fish, the
RT   spiny dog-fish (Squalus acanthias).";
RL   C. R. Acad. Sci., D, Sci. Nat. 274:313-316(1972).
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   InterPro; IPR000981; Neurhyp_horm.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Amidation.
FT   DISULFID      1      6
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  996 MW;  17F8376EB444404B CRC64;

  Query Match             28.9%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   28.6%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 YLSGADL 7
              |::   |
Db          2 YINNCPL 8


RESULT 13
OXYT_RAJCL
ID   OXYT_RAJCL     STANDARD;      PRT;     9 AA.
AC   P42994;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Glumitocin.
OS   Raja clavata (Thornback ray).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;
OC   Rajiformes; Rajidae; Raja.
OX   NCBI_TaxID=7781;
RN   [1]
```

```
RP   SEQUENCE.
RX   MEDLINE=66123415; PubMed=5880565;
RA   Acher R., Chauvet J., Chauvet M.-T., Crepy D.;
RT   "Phylogeny of neurophyophyseal peptides: isolation of a new hormone,
RT   glumitocin (Ser 4-Gln 8-ocytocin) present in a cartilaginous fish,
RT   the ray (Raia clavata).";
RL   Biochim. Biophys. Acta 107:393-396(1965).
CC   -!- FUNCTION: ANTIDIURETIC HORMONE.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   InterPro; IPR000981; Neurhyp_horm.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Amidation.
FT   DISULFID      1      6
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  984 MW;  17E9C76EB455B04B CRC64;

  Query Match             28.9%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLS 3
              |:|
Db          2 YIS 4


RESULT 14
PPH2_LYCES
ID   PPH2_LYCES     STANDARD;      PRT;     7 AA.
AC   P83379;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Purple acid phosphatase isozyme LeSAP2 (EC 3.1.3.2) (Fragment).
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE, CATALYTIC ACTIVITY, SUBUNIT, SUBCELLULAR LOCATION, AND
RP   GLYCOSYLATION.
RC   STRAIN=cv. Moneymaker; TISSUE=Seed;
RX   MEDLINE=22361242; PubMed=12473124;
RA   Bozzo G.G., Raghothama K.G., Plaxton W.C.;
RT   "Purification and characterization of two secreted purple acid
RT   phosphatase isozymes from phosphate-starved tomato (Lycopersicon
RT   esculentum) cell cultures.";
RL   Eur. J. Biochem. 269:6278-6286(2002).
CC   -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC       alcohol + phosphate.
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: Glycosylated.
CC   -!- MISCELLANEOUS: In L.esculentum there are at least two isozymes of
CC       purple acid phosphatase.
KW   Hydrolase; Glycoprotein.
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  810 MW;  672AA862C9C729A0 CRC64;

  Query Match             26.7%;  Score 12;  DB 1;  Length 7;
  Best Local Similarity   42.9%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 YLSGADL 7
              :|   ||
Db          1 FLFVGDL 7


RESULT 15
CCKN_MACEU
ID   CCKN_MACEU     STANDARD;      PRT;     8 AA.
AC   P30369;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Cholecystokinin (CCK).
GN   CCK.
OS   Macropus eugenii (Tammar wallaby), and
OS   Dasyurus viverrinus (Southeastern quoll).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX   NCBI_TaxID=9315, 9279;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=M.eugenii, and D.viverrinus;
RC   TISSUE=Brain;
RX   MEDLINE=88234141; PubMed=3375140;
RA   Fan Z.W., Eng J., Shaw G., Yalow R.S.;
RT   "Cholecystokinin octapeptide purified from brains of Australian
RT   marsupials.";
RL   Peptides 9:429-431(1988).
CC   -!- FUNCTION: THIS PEPTIDE HORMONE INDUCES GALL BLADDER CONTRACTION
CC       AND THE RELEASE OF PANCREATIC ENZYMES IN THE GUT. ITS FUNCTION
CC       IN THE BRAIN IS NOT CLEAR.
CC   -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
DR   PIR; A43001; A43001.
DR   PIR; PQ0012; PQ0012.
DR   InterPro; IPR001651; Gastrin.
DR   PROSITE; PS00259; GASTRIN; 1.
KW   Amidation; Sulfation; Hormone.
FT   MOD_RES       2      2       SULFATION.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  1064 MW;  DDCAA68378768B5A CRC64;

  Query Match             26.7%;  Score 12;  DB 1;  Length 8;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 YLSGAD 6
              |:   |
Db          2 YMGWMD 7


Search completed: January 12, 2004, 14:29:02
Job time : 7.25 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:        January 12, 2004, 14:24:44 ; Search time 25.75 Seconds
                                          (without alignments)
                                          90.193 Million cell updates/sec

Title:         US-09-529-121A-2
Perfect score: 45
Sequence:      1 YLSGADLNL 9

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters:      775

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_23:*
               1:  sp_archea:*
               2:  sp_bacteria:*
               3:  sp_fungi:*
               4:  sp_human:*
               5:  sp_invertebrate:*
               6:  sp_mammal:*
               7:  sp_mhc:*
               8:  sp_organelle:*
               9:  sp_phage:*
               10:  sp_plant:*
               11:  sp_rodent:*
               12:  sp_virus:*
               13:  sp_vertebrate:*
               14:  sp_unclassified:*
               15:  sp_rvirus:*
               16:  sp_bacteriap:*
               17:  sp_archeap:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                    SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                      Description
-------------------------------------------------------------------------
    1      20    44.4      8  3  Q9HDS4                  Q9hds4 aspergillus
    2      17    37.8      7 11  Q63480                   Q63480 rattus norv
    3      17    37.8      8  6  Q8WNS1                  Q8wns1 bos taurus
    4      16    35.6      8 13  P82079                   P82079 limnodynast
    5      15    33.3      8  3  P87225                  P87225 saccharomyc
    6      15    33.3      8  4  Q9UMC7                  Q9umc7 homo sapien
    7      14    31.1      8  2  Q9X3K1                  Q9x3k1 prochloroco
    8      14    31.1      8  5  Q94695                  Q94695 physarum po
    9      14    31.1      8 13  Q9PS69                   Q9ps69 gallus gall
   10      14    31.1      9  1  Q50832                  Q50832 methanococc
   11      14    31.1      9  4  Q16220                  Q16220 homo sapien
   12      14    31.1      9  4  O95953                  O95953 homo sapien
   13      14    31.1      9  6  Q9TRS0                  Q9trs0 oryctolagus
   14      14    31.1      9 11  O35953                   O35953 mus musculu
   15      13    28.9      7  2  Q8KMS3                  Q8kms3 klebsiella
   16      13    28.9      7 11  O55184                   O55184 rattus norv
   17      13    28.9      9  2  P83222                  P83222 streptomyce
   18      13    28.9      9  5  Q9TWD6                  Q9twd6 leptinotars
   19      13    28.9      9  6  Q9TRW2                  Q9trw2 oryctolagus
   20      13    28.9      9 13  Q8JJ12                   Q8jj12 carassius a
   21      13    28.9      9 13  Q8JJ18                   Q8jj18 danio aequi
   22      13    28.9      9 13  Q8JJ14                   Q8jj14 danio frank
   23      13    28.9      9 13  Q8JJ10                   Q8jj10 notropis ch
   24      13    28.9      9 13  Q8JJ16                   Q8jj16 danio albol
   25      12    26.7      8  2  P77556                  P77556 escherichia
   26      12    26.7      8  4  Q15901                  Q15901 homo sapien
   27      12    26.7      8  4  Q9UMH9                  Q9umh9 homo sapien
   28      12    26.7      8  4  Q8IUB8                  Q8iub8 homo sapien
   29      12    26.7      8  7  Q95213                  Q95213 oryctolagus
   30      12    26.7      8 13  Q91098                   Q91098 manorina me
   31      12    26.7      8 13  Q90498                   Q90498 erythrura g
   32      12    26.7      9  2  Q9R635                  Q9r635 chlamydia t
   33      12    26.7      9  2  Q57328                  Q57328 aeromonas s
   34      12    26.7      9  2  Q44377                  Q44377 aeromonas t
   35      12    26.7      9  2  Q44468                  Q44468 aeromonas v
   36      12    26.7      9  2  Q43928                  Q43928 aeromonas p
   37      12    26.7      9  2  Q44001                  Q44001 aeromonas e
   38      12    26.7      9  4  Q9H522                  Q9h522 homo sapien
   39      12    26.7      9  4  Q9UCN5                  Q9ucn5 homo sapien
   40      12    26.7      9  5  Q27396                  Q27396 babesia bov
   41      12    26.7      9  5  O96417                  O96417 drosophila
   42      12    26.7      9  6  Q28112                  Q28112 bos taurus
   43      12    26.7      9  8  Q94NB1                  Q94nb1 microcebus
   44      12    26.7      9  8  Q94NB2                  Q94nb2 microcebus
   45      12    26.7      9  8  Q94NA9                  Q94na9 daubentonia

                                 ALIGNMENTS


RESULT 1
Q9HDS4
ID   Q9HDS4     PRELIMINARY;      PRT;     8 AA.
AC   Q9HDS4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   TrpC polyprotein (Fragment).
GN   TRPC.
OS   Aspergillus flavus.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX   NCBI_TaxID=5059;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A55;
RA   Geiser D.M., Dorner J.W., Horn B.W., Taylor J.W.;
RT   "The phylogenetics of mycotoxin and sclerotium production in
RT   Aspergillus flavus and Aspergillus oryzae.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF261861; AAG16135.1; -.
KW   Polyprotein.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  807 MW;  F3B2C72AB5B87DD6 CRC64;

  Query Match             44.4%;  Score 20;  DB 3;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSGADL 7
              ::|:||
Db          1 MAGSDL 6


RESULT 2
Q63480
ID   Q63480     PRELIMINARY;      PRT;     7 AA.
AC   Q63480;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
```

```
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TR4-NS orphan receptor (Fragment).
GN   TR4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96198747; PubMed=8612486;
RA   Yoshikawa T., Makino S., Gao X.M., Xing G.Q., Chuang D.M.,
RA   Detera-Wadleigh S.D.;
RT   "Splice variants of rat TR4 orphan receptor: differential expression
RT   of novel sequences in the 5'-untranslated region and C-terminal
RT   domain.";
RL   Endocrinology 137:1562-1571(1996).
DR   EMBL; U59125; AAB02827.1; -.
KW   Receptor.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  758 MW;  672AA87864005350 CRC64;

  Query Match             37.8%;  Score 17;  DB 11;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 LSGADL 7
              : | ||
Db          2 IRGGDL 7


RESULT 3
Q8WNS1
ID   Q8WNS1      PRELIMINARY;      PRT;     8 AA.
AC   Q8WNS1;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   X-linked zinc finger protein (Fragment).
GN   ZFX.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Poloumienko A., Blecher S.;
RT   "Comparison between intron-exon structures in ZFX and ZFY genes.";
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF045782; AAL58190.1; -.
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  904 MW;  DF1DC2C4472AAB1A CRC64;

  Query Match             37.8%;  Score 17;  DB 6;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DLNL 9
              |||:
Db          3 DLNV 6


RESULT 4
P82079
ID   P82079      PRELIMINARY;      PRT;     8 AA.
AC   P82079;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DYNASTIN 1.
OS   Limnodynastes interioris (Giant banjo frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Limnodynastinae; Limnodynastes.
OX   NCBI_TaxID=30362;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=TIBIAL GLAND;
RA   Raftery M.J., Bradford A.M., Bowie J.H., Wallace J.C., Tyler M.J.;
RT   "Peptides from Australian frogs. The structure of the dynastins from
RT   the banjo frogs Limnodynastes interioras, Limnodynastes dumerilii and
RT   Limnodynastes terraereginae.";
RL   Aust. J. Chem. 46:833-842(1993).
CC   -!- MASS SPECTROMETRY: MW=729; METHOD=FAB.
KW   Amphibian skin.
SQ   SEQUENCE   8 AA;  729 MW;  7C28772865B72728 CRC64;

  Query Match             35.6%;  Score 16;  DB 13;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 8.3e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 LSGADL 7
              |||  |
Db          3 LSGLGL 8


RESULT 5
P87225
ID   P87225      PRELIMINARY;      PRT;     8 AA.
AC   P87225;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   GIN11 protein (Fragment).
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wedler H., Wedler E., Scharfe M., Wambutt R.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   MIPS;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z73169; CAA97518.2; -.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  1019 MW;  4E21A9C449D5B73B CRC64;

  Query Match             33.3%;  Score 15;  DB 3;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLS 3
              |||
Db          1 YLS 3


RESULT 6
Q9UMC7
ID   Q9UMC7      PRELIMINARY;      PRT;     8 AA.
AC   Q9UMC7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   SHMT protein (Fragment).
GN   SHMT.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Chave K.J., Snell K., Sanders P.G.;
RT   "Isolation and characterisation of human genomic sequences encoding
RT   cytosolic serine hydroxymethyltransferase.";
RL   Biochem. Soc. Trans. 25:53-53(1997).
DR   EMBL; Y14492; CAB54844.1; -.
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  868 MW;  7C205721E44AB5B8 CRC64;

  Query Match             33.3%;  Score 15;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GADLNL 9
              |:| :|
Db          1 GSDNHL 6


RESULT 7
Q9X3K1
ID   Q9X3K1      PRELIMINARY;      PRT;     8 AA.
AC   Q9X3K1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   Cytochrome b (Fragment).
GN   PETB.
OS   Prochlorococcus sp.
OC   Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC   Prochlorococcus.
OX   NCBI_TaxID=1220;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Urbach E., Chisholm S.W.;
RT   "Genetic diversity in Prochlorococcus populations flow cytometrically
RT   sorted from the Sargasso Sea and Gulf Stream.";
RL   Limnol. Oceanog. 43:1615-1630(1998).
DR   EMBL; AF070193; AAD23233.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  799 MW;  10376865B72866D3 CRC64;

  Query Match             31.1%;  Score 14;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |||
Db          4 LSG 6


RESULT 8
Q94695
ID   Q94695      PRELIMINARY;      PRT;     8 AA.
AC   Q94695;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   Actin (Fragment).
GN   ARDC.
OS   Physarum polycephalum (Slime mold).
OC   Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physarida;
OC   Physarum.
OX   NCBI_TaxID=5791;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96182101; PubMed=8622700;
RA   Benard M., Lagnel C., Pallotta D., Pierron G.;
RT   "Mapping of a replication origin within the promoter region of two
RT   unlinked, abundantly transcribed actin genes of Physarum
RT   polycephalum.";
RL   Mol. Cell. Biol. 16:968-976(1996).
DR   EMBL; M73459; AAB03706.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  878 MW;  F4C6C2CAAB187B16 CRC64;

  Query Match             31.1%;  Score 14;  DB 5;  Length 8;
  Best Local Similarity   33.3%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 LSGADL 7
              : | |:
Db          1 MEGEDV 6


RESULT 9
Q9PS69
ID   Q9PS69      PRELIMINARY;      PRT;     8 AA.
AC   Q9PS69;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   LOW density lipoprotein receptor-related protein (Fragment).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92011685; PubMed=1918027;
RA   Stifani S., Barber D.L., Aebersold R., Steyrer E., Shen X., Nimpf J.,
RA   Schneider W.J.;
RT   "The laying hen expresses two different low density lipoprotein
RT   receptor-related proteins.";
RL   J. Biol. Chem. 266:19079-19087(1991).
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  846 MW;  C007272DD865BAAA CRC64;

  Query Match             31.1%;  Score 14;  DB 13;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 SGA 5
              |||
Db          3 SGA 5


RESULT 10
Q50832
ID   Q50832      PRELIMINARY;      PRT;     9 AA.
AC   Q50832;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1996 (TrEMBLrel. 01, Last annotation update)
DE   Intergenic AT-rich DNA sequence (Fragment).
OS   Methanococcus voltae.
OC   Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC   Methanococcaceae; Methanococcus.
OX   NCBI_TaxID=2188;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85230552; PubMed=4006907;
RA   Bollschweiler C., Kuehn R., Klein A.;
RT   "Non-repetitive AT-rich sequences are found in intergenic regions of
RT   Methanococcus voltae DNA.";
RL   EMBO J. 4:805-809(1985).
DR   EMBL; X02518; CAA26355.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1087 MW;  99ED005DC404405A CRC64;

  Query Match             31.1%;  Score 14;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DLN 8
```

```
               |:|
Db           2 DIN 4


RESULT 11
Q16220
ID   Q16220      PRELIMINARY;      PRT;     9 AA.
AC   Q16220;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   HGRP protein (Fragment).
GN   HGRP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94320083; PubMed=8044796;
RA   Nagalla S.R., Spindel E.R.;
RT   "Functional analysis of the 5'-flanking region of the human gastrin-
RT   releasing peptide gene in small cell lung carcinoma cell lines.";
RL   Cancer Res. 54:4461-4467(1994).
DR   EMBL; S73265; AAD14116.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1070 MW;  77FE37672B040864 CRC64;

  Query Match            31.1%;  Score 14;  DB 4;  Length 9;
  Best Local Similarity   37.5%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy           2 LSGADLNL 9
               : | :| |
Db           1 MRGRELPL 8


RESULT 12
O95953
ID   O95953      PRELIMINARY;      PRT;     9 AA.
AC   O95953;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Galactocerebrosidase (EC 3.2.1.46) (Fragment).
GN   GALC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Lulli L., Torchiana E., Finocchiaro G.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U77631; AAD15626.1; -.
KW   Glycosidase; Hydrolase.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1069 MW;  374E2AADC2C699C8 CRC64;

  Query Match            31.1%;  Score 14;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           5 ADL 7
               |||
Db           6 ADL 8


RESULT 13
Q9TRS0
ID   Q9TRS0      PRELIMINARY;      PRT;     9 AA.
AC   Q9TRS0;
```

```
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Calcyclin-associated protein, CAP50=CA2+/phospholipid-binding protein
DE   L-7 fragment (Fragment).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92250478; PubMed=1533622;
RA   Tokumitsu H., Mizutani A., Minami H., Kobayashi R., Hidaka H.;
RT   "A calcyclin-associated protein is a newly identified member of the
RT   Ca2+/phospholipid-binding proteins, annexin family.";
RL   J. Biol. Chem. 267:8919-8924(1992).
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1010 MW;  64E419C44865B72B CRC64;

  Query Match            31.1%;  Score 14;  DB 6;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           2 LSG 4
               |||
Db           3 LSG 5


RESULT 14
O35953
ID   O35953      PRELIMINARY;      PRT;     9 AA.
AC   O35953;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Truncated voltage-gated sodium channel alpha subunit (Fragment).
GN   SCN8A.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RIII;
RX   MEDLINE=97442476; PubMed=9295353;
RA   Plummer N.W., McBurney M.W., Meisler M.H.;
RT   "Alternative splicing of the sodium channel SCN8A predicts a truncated
RT   two-domain protein in fetal brain and non-neuronal cells.";
RL   J. Biol. Chem. 272:24008-24015(1997).
DR   EMBL; U97672; AAB80914.1; -.
DR   MGD; MGI:103169; Scn8a.
KW   Ionic channel.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  898 MW;  22D92865B735B737 CRC64;

  Query Match            31.1%;  Score 14;  DB 11;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           2 LSG 4
               |||
Db           5 LSG 7


RESULT 15
Q8KMS3
ID   Q8KMS3      PRELIMINARY;      PRT;     7 AA.
AC   Q8KMS3;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Putative MerR2 protein.
```

```
GN   MERR2.
OS   Klebsiella sp. LS13-39.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Klebsiella.
OX   NCBI_TaxID=143776;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LS13-39;
RX   MEDLINE=21604134; PubMed=11763242;
RA   Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,
RA   Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,
RA   Yurieva O.V., Nikiforov V.G.;
RT   "Mercury resistance transposons of Gram-negative environmental
RT   bacteria and their classification.";
RL   Res. Microbiol. 152:811-822(2001).
DR   EMBL; AJ302776; CAC82975.1; -.
SQ   SEQUENCE   7 AA;  608 MW;  6DC1B5BDD87DD6F0 CRC64;

  Query Match             28.9%;  Score 13;  DB 2;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSGA 5
              ::||
Db          1 MAGA 4


Search completed: January 12, 2004, 14:30:59
Job time : 27.75 secs
```

```
                    GenCore version 5.1.6
             Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         January 12, 2004, 14:25:20 ; Search time 10.25 Seconds
                                             (without alignments)
                                             84.441 Million cell updates/sec

Title:          US-09-529-121A-3
Perfect score:  45
Sequence:       1 YLSGADINL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters:      789

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_76:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                    Description
 ---------------------------------------------------------------------------
    1      16    35.6      7  2  PT0542                T-cell receptor be
    2      16    35.6      8  2  A21440                variant surface gl
    3      16    35.6      8  2  JS0318                leucokinin VIII -
    4      16    35.6      9  2  A61364                isotocin - common
    5      16    35.6      9  4  I57650                hemoglobin alpha c
    6      15    33.3      5  2  PT0540                T-cell receptor be
    7      15    33.3      6  2  PT0726                T-cell receptor be
    8      15    33.3      7  2  PT0526                T-cell receptor be
    9      15    33.3      7  2  PT0676                T-cell receptor be
   10      15    33.3      8  2  T13818                cytochrome oxidase
   11      15    33.3      9  2  F41978                calliFMRFamide 6 -
   12      15    33.3      9  2  PT0288                Ig heavy chain CRD
   13      15    33.3      9  2  G41946                T-cell receptor ga
   14      14    31.1      5  2  PT0679                T-cell receptor be
   15      14    31.1      6  2  PT0605                T-cell receptor be
   16      14    31.1      6  2  PT0593                T-cell receptor be
   17      14    31.1      7  2  PT0654                T-cell receptor be
   18      14    31.1      7  2  PT0722                T-cell receptor be
   19      14    31.1      7  2  PC2370                probable H+-transp
   20      14    31.1      8  2  S65647                2-hydroxyglutaryl-
   21      14    31.1      8  2  PN0043                phosphatidylethano
   22      14    31.1      8  2  PT0557                T-cell receptor be
   23      14    31.1      9  1  AKLQIM                locustamyoinhibiti
   24      14    31.1      9  2  A57444                neuropeptide Grb-A
   25      14    31.1      9  2  B57444                neuropeptide Grb-A
   26      14    31.1      9  2  C57444                neuropeptide Grb-A
   27      14    31.1      9  2  PT0268                Ig heavy chain CRD
   28      13    28.9      4  2  S43959                Ig mu chain V regi
   29      13    28.9      7  2  I50210                gene c-rel protein
   30      13    28.9      7  2  A58718                carnocin UI49 - Ca
   31      13    28.9      8  2  A41117                acetylcholinestera
   32      13    28.9      8  2  PT0547                T-cell receptor be
   33      13    28.9      9  2  QDRB                  delta sleep-induci
   34      13    28.9      9  2  C41170                photosystem II pro
   35      13    28.9      9  2  A61386                macrophage inhibit
   36      13    28.9      9  2  PH0935                T-cell receptor be
   37      13    28.9      9  2  PH0918                T-cell receptor be
   38      12    26.7      4  2  A26209                protein-glutamine
   39      12    26.7      5  2  S62883                seminal plasma pro
   40      12    26.7      6  2  I51434                H4 histone - Afric
   41      12    26.7      7  2  S16364                opacity protein P.
   42      12    26.7      7  2  B35890                RNA-directed DNA p
   43      12    26.7      7  2  S20446                elastase - Pseudom
   44      12    26.7      7  2  A34818                vicilin 72K chain
   45      12    26.7      7  2  S29735                polyphosphate-gluc

                                   ALIGNMENTS

RESULT 1
PT0542
T-cell receptor beta chain V-D-J region (126-1BA) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0542
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0542
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: day 18 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             35.6%;  Score 16;  DB 2;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              || |
Db          2 SGGD 5


RESULT 2
A21440
variant surface glycoprotein pSLc1 - Trypanosoma brucei (fragment)
C;Species: Trypanosoma brucei
C;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 20-Mar-1998
C;Accession: A21440
R;Parsons, M.; Nelson, R.G.; Watkins, K.P.; Agabian, N.
Cell 38, 309-316, 1984
A;Title: Trypanosome mRNAs share a common 5' spliced leader sequence.
A;Reference number: A90853; MUID:84282716; PMID:6088073
A;Accession: A21440
A;Molecule type: mRNA
A;Residues: 1-8 <PAR>
A;Cross-references: GB:K02195; NID:g162150; PID:g162151
C;Keywords: glycoprotein

  Query Match             35.6%;  Score 16;  DB 2;  Length 8;
  Best Local Similarity   33.3%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 LSGADI 7
              :|| ::
Db          1 MSGKEV 6


RESULT 3
```

```
JS0318
leucokinin VIII - Madeira cockroach
C;Species: Leucophaea maderae (Madeira cockroach)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Jun-2000
C;Accession: JS0318
R;Holman, G.M.; Cook, B.J.; Nachman, R.J.
Comp. Biochem. Physiol. C 88, 31-34, 1987
A;Title: Isolation, primary structure and synthesis of leucokinins VII and VIII: the fir
A;Reference number: JS0317
A;Accession: JS0318
A;Molecule type: protein
A;Residues: 1-8 <HOL>
C;Comment: Leucokinins, a family of cephalomyotropic peptides, stimulate contractile act
C;Keywords: amidated carboxyl end; cephalomyotropic peptide
F;8/Modified site: amidated carboxyl end (Gly) #status experimental

  Query Match             35.6%;  Score 16;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GAD 6
              |||
Db          1 GAD 3


RESULT 4
A61364
isotocin - common carp
C;Species: Cyprinus carpio (common carp)
C;Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 20-Jun-2000
C;Accession: A61364
R;Acher, R.; Chauvet, J.; Chauvet, M.T.; Crepy, D.
Comp. Biochem. Physiol. A 14, 245-254, 1965
A;Title: Caracterisation des hormones neurohypophysaires d'un poisson osseux d'eau douce
A;Reference number: A61364
A;Accession: A61364
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <ACH>
C;Superfamily: oxytocin-neurophysin
C;Keywords: amidated carboxyl end; neuropeptide; posterior pituitary
F;9/Modified site: amidated carboxyl end (Gly) #status experimental

  Query Match             35.6%;  Score 16;  DB 2;  Length 9;
  Best Local Similarity   42.9%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 YLSGADI 7
                |:|  |
Db          2 YISNCPI 8


RESULT 5
I57650
hemoglobin alpha chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 31-Jul-1997 #text_change 20-Apr-2000
C;Accession: I57650
R;Whitelaw, E.; Hogben, P.; Hanscombe, O.; Proudfoot, N.J.
Mol. Cell. Biol. 9, 241-251, 1989
A;Title: Transcriptional promiscuity of the human alpha-globin gene.
A;Reference number: I57650; MUID:89181576; PMID:2538719
A;Accession: I57650
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-9 <WHI>
A;Cross-references: GB:M23454; NID:g340922; PIDN:AAA52629.1; PID:g553329
A;Note: engineered sequence; this sequence was not determined in this report
C;Genetics:
A;Gene: GDB:HBA1
A;Cross-references: GDB:119293
A;Map position: 16p13.3-16p13.3
```

```
  Query Match             35.6%;  Score 16;  DB 4;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 2.8e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 LSGAD 6
              || ||
Db          3 LSPAD 7


RESULT 6
PT0540
T-cell receptor beta chain V-D-J region (126-1L) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0540
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0540
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: day 18 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             33.3%;  Score 15;  DB 2;  Length 5;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              || |
Db          2 SGED 5


RESULT 7
PT0726
T-cell receptor beta chain V-D-J region (161-2D) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0726
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0726
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-6 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             33.3%;  Score 15;  DB 2;  Length 6;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              || |
Db          2 SGED 5


RESULT 8
PT0526
T-cell receptor beta chain V-D-J region (100-4E) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0526
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0526
A;Status: translation not shown
```

```
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: adult thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             33.3%;  Score 15;  DB 2;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              || |
Db          2 SGED 5


RESULT 9
PT0676
T-cell receptor beta chain V-D-J region (140-1AL) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0676
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0676
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-7 <FEE>
A;Experimental source: day 18 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             33.3%;  Score 15;  DB 2;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              || |
Db          2 SGED 5


RESULT 10
T13818
cytochrome oxidase subunit I - Atlantic hagfish mitochondrion (fragment)
C;Species: mitochondrion Myxine glutinosa (Atlantic hagfish)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T13818
R;Delarbre, C.; Barriel, V.; Tillier, S.; Janvier, P.; Gachelin, G.
Mol. Biol. Evol. 14, 807-813, 1997
A;Title: The main features of the craniate mitochondrial DNA between the ND1 and the COI
A;Reference number: Z17775; MUID:97398704; PMID:9254918
A;Accession: T13818
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-8 <DEL>
A;Cross-references: EMBL:Y09527; NID:g2340019; PIDN:CAA70718.1; PID:g2340022
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion

  Query Match             33.3%;  Score 15;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLS 3
              |||
Db          2 YLS 4


RESULT 11
F41978
calliFMRFamide 6 - bluebottle fly (Calliphora vomitoria)
C;Species: Calliphora vomitoria
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
C;Accession: F41978
R;Duve, H.; Johnsen, A.H.; Sewell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, J.F.; Thorpe,
Proc. Natl. Acad. Sci. U.S.A. 89, 2326-2330, 1992
A;Title: Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH-2 neuropeptides (desig
A;Reference number: A41978; MUID:92196111; PMID:1549595
A;Accession: F41978
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <DUV>
C;Keywords: amidated carboxyl end; neuropeptide
F;9/Modified site: amidated carboxyl end (Phe) #status experimental

  Query Match             33.3%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              || |
Db          2 SGQD 5


RESULT 12
PT0288
Ig heavy chain CRD3 region (clone 4-106) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0288
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and jo
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0288
A;Molecule type: DNA
A;Residues: 1-9 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

  Query Match             33.3%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSG 4
              | ||
Db          5 YSSG 8


RESULT 13
G41946
T-cell receptor gamma chain (2t.23) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C;Accession: G41946
R;Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A;Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ger
A;Reference number: A41946; MUID:92049316; PMID:1658619
A;Accession: G41946
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-9 <WHE>
C;Keywords: T-cell receptor

  Query Match             33.3%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSG 4
              | ||
Db          5 YSSG 8


RESULT 14
```

```
PT0679
T-cell receptor beta chain V-D-J region - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0679; PT0708
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0679
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-5 <FEE>
A;Experimental source: day 18 fetal thymus, strain BALB/c, 154-2J
A;Accession: PT0708
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-5 <FE2>
A;Experimental source: newborn thymus, strain BALB/c, 161-2B
C;Keywords: T-cell receptor

  Query Match            31.1%;  Score 14;  DB 2;  Length 5;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              || |
Db          2 SGDD 5


RESULT 15
PT0605
T-cell receptor beta chain V-D-J region (120-1L) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0605
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0605
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match            31.1%;  Score 14;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 SGA 5
              |||
Db          2 SGA 4


Search completed: January 12, 2004, 14:31:52
Job time : 10.25 secs
```

```
                      GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:        January 12, 2004, 14:21:34 ; Search time 6.25 Seconds
                                            (without alignments)
                                            67.718 Million cell updates/sec

Title:         US-09-529-121A-3
Perfect score: 45
Sequence:      1 YLSGADINL 9

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters:      251

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_41:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                  %
Result           Query
  No.   Score    Match Length DB  ID                      Description
 -----------------------------------------------------------------------------
   1      17      37.8      8  1  CPD1_ENTFA              P13269 enterococcu
   2      16      35.6      8  1  LCK8_LEUMA              P19990 leucophaea
   3      16      35.6      9  1  ISOT_CYPCA              P42993 cyprinus ca
   4      15      33.3      8  1  PLP_BRANA               P81707 brassica na
   5      15      33.3      9  1  FAR6_CALVO              P41861 calliphora
   6      14      31.1      8  1  ALL5_CYDPO              P82156 cydia pomon
   7      14      31.1      9  1  LMIP_LOCMI              P31799 locusta mig
   8      14      31.1      9  1  PTSP_BOMMO              P82003 bombyx mori
   9      13      28.9      7  1  LANC_CARUI              P36960 carnobacter
  10      13      28.9      8  1  FAR2_MACRS              P83275 macrobrachi
  11      13      28.9      9  1  DSIP_RABIT              P01158 oryctolagus
  12      13      28.9      9  1  FAR2_PANRE              P41873 panagrellus
  13      13      28.9      9  1  OXYT_RAJCL              P42994 raja clavat
  14      12      26.7      8  1  ACT_CARMA               P80709 carcinus ma
  15      12      26.7      8  1  CCKN_MACEU              P30369 macropus eu
  16      12      26.7      9  1  FARD_CALVO              P41868 calliphora
  17      12      26.7      9  1  OXYA_SCYCA              P42996 scyliorhinu
  18      12      26.7      9  1  OXYT_BUFRE              P42995 bufo regula
  19      12      26.7      9  1  PGLR_DIAAB              P81179 diaprepes a
  20      12      26.7      9  1  UPA6_HUMAN              P30092 homo sapien
  21      11      24.4      4  1  FAR3_HIRME              P42562 hirudo medi
  22      11      24.4      5  1  PRCT_PERAM              P01373 periplaneta
  23      11      24.4      7  1  FAR2_ASCSU              P31890 ascaris suu
  24      11      24.4      7  1  GFRP_MOUSE              P99025 mus musculu
  25      11      24.4      8  1  CAD1_ENTFA              P13268 enterococcu
  26      11      24.4      8  1  FAR8_CALVO              P41863 calliphora
  27      11      24.4      9  1  D1_NEPNO                P24816 nephrops no
  28      11      24.4      9  1  FAR5_CALVO              P41860 calliphora
  29      11      24.4      9  1  FAR7_CALVO              P41862 calliphora
  30      11      24.4      9  1  MOSF_CLYJA              P19853 clypeaster
  31      11      24.4      9  1  OXYA_SQUAC              P42999 squalus aca
  32      11      24.4      9  1  OXYV_SQUAC              P43000 squalus aca
  33      10      22.2      4  1  ACH1_ACHFU              P35904 achatina fu
  34      10      22.2      5  1  UXA4_CHLTR              P38005 chlamydia t
  35      10      22.2      6  1  CIP2_MYTED              P13737 mytilus edu
  36      10      22.2      6  1  TMOF_SARBU              P41495 sarcophaga
  37      10      22.2      6  1  UN06_CLOPA              P81351 clostridium
  38      10      22.2      7  1  ALL2_CARMA              P81805 carcinus ma
  39      10      22.2      7  1  ALL3_CARMA              P81806 carcinus ma
  40      10      22.2      7  1  ALL4_CARMA              P81807 carcinus ma
  41      10      22.2      7  1  ALL5_CARMA              P81808 carcinus ma
  42      10      22.2      7  1  ALL7_CYDPO              P82158 cydia pomon
  43      10      22.2      7  1  PPH2_LYCES              P83379 lycopersico
  44      10      22.2      7  1  UN06_PINPS              P81675 pinus pinas
  45      10      22.2      8  1  AL12_CARMA              P81815 carcinus ma

                              ALIGNMENTS


RESULT 1
CPD1_ENTFA
ID   CPD1_ENTFA     STANDARD;      PRT;     8 AA.
AC   P13269;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Sex pheromone cPD1.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=85040388; PubMed=6436978;
RA   Suzuki A., Mori M., Sagakami Y., Isogai A., Fujino M., Kitada C.,
RA   Craig R.A., Clewell D.B.;
RT   "Isolation and structure of bacterial sex pheromone, cPD1.";
RL   Science 226:849-850(1984).
CC   -!- FUNCTION: CPD1 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC       BACTERIOCIN PLASMID PPD1.
KW   Pheromone.
SQ   SEQUENCE   8 AA;  913 MW;  8665B729C682C729 CRC64;

  Query Match            37.8%;  Score 17;  DB 1;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSG 4
              :|||
Db          5 FLSG 8


RESULT 2
LCK8_LEUMA
ID   LCK8_LEUMA     STANDARD;      PRT;     8 AA.
AC   P19990;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Leucokinin VIII (L-VIII).
OS   Leucophaea maderae (Madeira cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
OC   Blaberidae; Leucophaea.
OX   NCBI_TaxID=6988;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Head;
RA   Holman G.M., Cook B.J., Nachman R.J.;
RT   "Isolation, primary structure and synthesis of leucokinins VII and
RT   VIII: the final members of this new family of cephalomyotropic
RT   peptides isolated from head extracts of Leucophaea maderae.";
RL   Comp. Biochem. Physiol. 88C:31-34(1987).
CC   -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC       ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC   -!- SIMILARITY: TO THE OTHER LEUCOKININS.
```

```
DR   PIR; JS0318; JS0318.
KW   Neuropeptide; Amidation.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  902 MW;  736365AB59CAADD8 CRC64;

  Query Match             35.6%;  Score 16;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GAD 6
              |||
Db          1 GAD 3


RESULT 3
ISOT_CYPCA
ID   ISOT_CYPCA     STANDARD;      PRT;     9 AA.
AC   P42993;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Isotocin.
OS   Cyprinus carpio (Common carp).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Cyprinus.
OX   NCBI_TaxID=7962;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Pituitary;
RA   Acher R., Chauvet J., Chauvet M.-T., Crepy D.;
RT   "Characterization of neurohypophyseal hormones from a fresh water bony
RT   fish, the carp (Cyprinus carpio). Comparison with hormones from sea
RT   water bony fishs.";
RL   Comp. Biochem. Physiol. 14:245-254(1965).
CC   -!- FUNCTION: ANTIDIURETIC HORMONE.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   PIR; A61364; A61364.
DR   InterPro; IPR000981; Neurhyp_horm.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Amidation.
FT   DISULFID      1      6
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  969 MW;  17FF476EB455B04B CRC64;

  Query Match             35.6%;  Score 16;  DB 1;  Length 9;
  Best Local Similarity   42.9%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 YLSGADI 7
              |:|   |
Db          2 YISNCPI 8


RESULT 4
PLP_BRANA
ID   PLP_BRANA      STANDARD;      PRT;     8 AA.
AC   P81707;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Plastidial lipid-associated protein (Fragment).
OS   Brassica napus (Rape).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Brassica.
OX   NCBI_TaxID=3708;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=CV. TOPAZ; TISSUE=Tapetum;
RX   MEDLINE=99349136; PubMed=10420651;
RA   Hernandez-Pinzon I., Ross J.H.E., Barnes K.A., Damant A.P.,
```

```
RA   Murphy D.J.;
RT   "Composition and role of tapetal lipid bodies in the biogenesis of the
RT   pollen coat of Brassica napus.";
RL   Planta 208:588-598(1999).
CC   -!- FUNCTION: MAY PLAY A STRUCTURAL ROLE IN THE ELAIOPLAST, A TAPETUM-
CC       SPECIFIC PLASTIDIAL LIPID ORGANELLE.
CC   -!- TISSUE SPECIFICITY: TAPETUM OF ANTHERS.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  989 MW;  9D7B1AA452CAA042 CRC64;

  Query Match             33.3%;  Score 15;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DIN 8
              |:|
Db          3 DVN 5


RESULT 5
FAR6_CALVO
ID   FAR6_CALVO     STANDARD;      PRT;     9 AA.
AC   P41861;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   CalliFMRFamide 6.
OS   Calliphora vomitoria (Blue blowfly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC   Calliphoridae; Calliphora.
OX   NCBI_TaxID=27454;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Thoracic ganglion;
RX   MEDLINE=92196111; PubMed=1549595;
RA   Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA   Rehfeld J.F., Thorpe A.;
RT   "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT   neuropeptides (designated calliFMRFamides) from the blowfly
RT   Calliphora vomitoria.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
DR   PIR; F41978; F41978.
KW   Neuropeptide; Amidation.
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1058 MW;  96D10699CAB6D865 CRC64;

  Query Match             33.3%;  Score 15;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              || |
Db          2 SGQD 5


RESULT 6
ALL5_CYDPO
ID   ALL5_CYDPO     STANDARD;      PRT;     8 AA.
AC   P82156;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Cydiastatin 5.
OS   Cydia pomonella (Codling moth).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Tortricoidea; Tortricidae; Olethreutinae; Cydia.
OX   NCBI_TaxID=82600;
RN   [1]
RP   SEQUENCE.
```

```
RC   TISSUE=Larva;
RX   MEDLINE=98054539; PubMed=9392829;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA   Davey M., East P.D., Thorpe A.;
RT   "Lepidopteran peptides of the allatostatin superfamily.";
RL   Peptides 18:1301-1309(1997).
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES        8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  898 MW;  922879CABB58640D CRC64;

  Query Match             31.1%;  Score 14;  DB 1;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 GADINL 9
              | |  |
Db          3 GYDFGL 8


RESULT 7
LMIP_LOCMI
ID   LMIP_LOCMI     STANDARD;      PRT;     9 AA.
AC   P31799;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-OCT-1993 (Rel. 27, Last annotation update)
DE   Locustamyoinhibiting peptide (LOM-MIP).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC   Acridoidea; Acrididae; Oedipodinae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92179466; PubMed=1796179;
RA   Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;
RT   "Isolation, identification and synthesis of locustamyoinhibiting
RT   peptide (LOM-MIP), a novel biologically active neuropeptide from
RT   Locusta migratoria.";
RL   Regul. Pept. 36:111-119(1991).
CC   -!- FUNCTION: SUPPRESSES SPONTANEOUS CONTRACTIONS OF THE HINDGUT AND
CC       OVIDUCT.
CC   -!- TISSUE SPECIFICITY: NEURONS LOCATED IN TWO VENTRAL CELL CLUSTERS
CC       IN THE SUBOESOPHAGEAL GANGLION.
DR   PIR; A60065; AKLQIM.
KW   Amidation; Neuropeptide.
FT   MOD_RES        9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1060 MW;  387D7DD4472AB6C3 CRC64;

  Query Match             31.1%;  Score 14;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DIN 8
              |:|
Db          4 DLN 6


RESULT 8
PTSP_BOMMO
ID   PTSP_BOMMO     STANDARD;      PRT;     9 AA.
AC   P82003;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Prothoracicostatic peptide (Bom-PTSP).
OS   Bombyx mori (Silk moth).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC   Bombycidae; Bombyx.
OX   NCBI_TaxID=7091;
RN   [1]
```

```
RP   SEQUENCE.
RC   STRAIN=C145 X N140; TISSUE=Brain;
RX   MEDLINE=20002634; PubMed=10531308;
RA   Hua Y.-J., Tanaka Y., Nakamura K., Sakakibara M., Nagata S.,
RA   Kataoka H.;
RT   "Identification of a prothoracicostatic peptide in the larval brain of
RT   the silkworm, Bombyx mori.";
RL   J. Biol. Chem. 274:31169-31173(1999).
RN   [2]
RP   ERRATUM.
RA   Hua Y.-J., Tanaka Y., Nakamura K., Sakakibara M., Nagata S.,
RA   Kataoka H.;
RL   J. Biol. Chem. 275:9892-9892(2000).
CC   -!- FUNCTION: Inhibits ecdysteroid biosynthesis in the prothoracic
CC       gland.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DEVELOPMENTAL STAGE: EARLY FIFTH INSTAR.
KW   Hormone; Amidation.
FT   MOD_RES        9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1090 MW;  3878C5B4472AB6C3 CRC64;

  Query Match             31.1%;  Score 14;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DIN 8
              |:|
Db          4 DLN 6


RESULT 9
LANC_CARUI
ID   LANC_CARUI     STANDARD;      PRT;     7 AA.
AC   P36960;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Lantibiotic carnocin UI49 (Fragment).
OS   Carnobacterium sp. (strain UI49).
OC   Bacteria; Firmicutes; Lactobacillales; Carnobacteriaceae;
OC   Carnobacterium.
OX   NCBI_TaxID=35782;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92321768; PubMed=1622206;
RA   Stoffels G., Nissen-Meyer J., Gudmundsdottir A., Sletten K., Holo H.,
RA   Nes I.F.;
RT   "Purification and characterization of a new bacteriocin isolated from
RT   a Carnobacterium sp.";
RL   Appl. Environ. Microbiol. 58:1417-1422(1992).
CC   -!- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC).
CC       ACTIVE ON GRAM-POSITIVE BACTERIA.
KW   Antibiotic; Bacteriocin; Lantibiotic.
FT   NON_TER        7      7
SQ   SEQUENCE   7 AA;  786 MW;  741776D05B05B810 CRC64;

  Query Match             28.9%;  Score 13;  DB 1;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GADI 7
              |::|
Db          1 GSEI 4


RESULT 10
FAR2_MACRS
ID   FAR2_MACRS     STANDARD;      PRT;     8 AA.
AC   P83275;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   FMRFamide-like neuropeptide FLP2 (ADKNFLRF-amide).
```

```
OS   Macrobrachium rosenbergii (Giant fresh water prawn).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea;
OC   Palaemonoidea; Palaemonidae; Macrobrachium.
OX   NCBI_TaxID=79674;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=Eyestalk;
RA   Sithigorngul P., Saraithongkum W., Jaideechoey S., Longyant S.,
RA   Sithigorngul W.;
RT   "Novel FMRFamide-like neuropeptides from the eyestalk of the giant
RT   freshwater prawn Macrobrachium rosenbergii.";
RL   Comp. Biochem. Physiol. 120B:587-595(1998).
CC   -!- MASS SPECTROMETRY: MW=1009.4; METHOD=MALDI.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
DR   GO; GO:0007218; P:neuropeptide signaling pathway; TAS.
KW   Neuropeptide; Amidation.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  1010 MW;  9CD40729C4433AAD CRC64;

  Query Match             28.9%;  Score 13;  DB 1;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 ADIN 8
              || |
Db          1 ADKN 4


RESULT 11
DSIP_RABIT
ID   DSIP_RABIT     STANDARD;      PRT;     9 AA.
AC   P01158;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Delta sleep-inducing peptide (DSIP).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=77185324; PubMed=862769;
RA   Monnier M., Dudler L., Gachter R., Maier P.F., Tobler H.J.,
RA   Schoenenberger G.A.;
RT   "The delta sleep inducing peptide (DSIP). Comparative properties of
RT   the original and synthetic nonapeptide.";
RL   Experientia 33:548-552(1977).
RN   [2]
RP   SEQUENCE, AND SYNTHESIS.
RX   MEDLINE=79054421; PubMed=568769;
RA   Schoenenberger G.A., Maier P.F., Tobler H.J., Wilson K., Monnier M.;
RT   "The delta EEG (sleep)-inducing peptide (DSIP). XI. Amino-acid
RT   analysis, sequence, synthesis and activity of the nonapeptide.";
RL   Pflugers Arch. 376:119-129(1978).
RN   [3]
RP   REVIEW.
RX   MEDLINE=87175129; PubMed=3550726;
RA   Graf M.V., Kastin A.J.;
RT   "Delta-sleep-inducing peptide (DSIP): an update.";
RL   Peptides 7:1165-1187(1986).
CC   -!- FUNCTION: WHEN INFUSED INTO THE MESODIENCEPHALIC VENTRICLE OF
CC       RECIPIENT RABBITS INDUCES SPINDLE AND DELTA EEG ACTIVITY AND
CC       REDUCED MOTOR ACTIVITIES.
CC   -!- MISCELLANEOUS: THIS PEPTIDE WAS OBTAINED FROM DIALYSATES OF
CC       OCCIPITAL VENOUS SINUS BLOOD FROM RABBITS KEPT ASLEEP BY ELECTRIC
CC       STIMULATION OF THE THALAMUS.
CC   -!- DATABASE: NAME=Protein Spotlight;
CC       NOTE=Issue 8 of March 2001;
CC       WWW="http://www.expasy.org/spotlight/articles/sptlt008.html".
DR   PIR; A01422; QDRB.
```

```
SQ   SEQUENCE   9 AA;  849 MW;  DDD365BDDAA8787D CRC64;

  Query Match             28.9%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAD 6
              :| |
Db          2 AGGD 5


RESULT 12
FAR2_PANRE
ID   FAR2_PANRE     STANDARD;      PRT;     9 AA.
AC   P41873;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFamide-like neuropeptide PF2 (SADPNFLRF-amide).
OS   Panagrellus redivivus.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
OC   Panagrolaimoidea; Panagrolaimidae; Panagrellus.
OX   NCBI_TaxID=6233;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=93027659; PubMed=1408999;
RA   Geary T.G., Price D.A., Bowman J.W., Winterrowd C.A., Mackenzie C.D.,
RA   Garrison R.D., Williams J.F., Friedman A.R.;
RT   "Two FMRFamide-like peptides from the free-living nematode
RT   Panagrellus redivivus.";
RL   Peptides 13:209-214(1992).
CC   -!- FUNCTION: MYOACTIVE.
CC   -!- TISSUE SPECIFICITY: NERVE CORDS AND PAIRED GROUPS OF CELLS LOCATED
CC       CAUDALLY TO THE BASE OF THE PHARYNX.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1066 MW;  DA0B0729C4576AAD CRC64;

  Query Match             28.9%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 ADIN 8
              || |
Db          2 ADPN 5


RESULT 13
OXYT_RAJCL
ID   OXYT_RAJCL     STANDARD;      PRT;     9 AA.
AC   P42994;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Glumitocin.
OS   Raja clavata (Thornback ray).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;
OC   Rajiformes; Rajidae; Raja.
OX   NCBI_TaxID=7781;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=66123415; PubMed=5880565;
RA   Acher R., Chauvet J., Chauvet M.-T., Crepy D.;
RT   "Phylogeny of neurophyophyseal peptides: isolation of a new hormone,
RT   glumitocin (Ser 4-Gln 8-ocytocin) present in a cartilaginous fish,
RT   the ray (Raia clavata).";
RL   Biochim. Biophys. Acta 107:393-396(1965).
CC   -!- FUNCTION: ANTIDIURETIC HORMONE.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   InterPro; IPR000981; Neurhyp_horm.
```

```
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Amidation.
FT   DISULFID      1      6
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  984 MW;  17E9C76EB455B04B CRC64;

  Query Match             28.9%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLS 3
              |:|
Db          2 YIS 4


RESULT 14
ACT_CARMA
ID   ACT_CARMA      STANDARD;      PRT;     8 AA.
AC   P80709;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Actin (Fragment).
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RA   Lachaise F., Somme G., Carpentier G., Granjeon E., Webster S.,
RA   Baghdassarian D.;
RT   "A transaldolase. An enzyme implicated in crab steroidogenesis.";
RL   Endocrine 5:23-32(1996).
CC   -!- FUNCTION: ACTINS ARE HIGHLY CONSERVED PROTEINS THAT ARE INVOLVED
CC       IN VARIOUS TYPES OF CELL MOTILITY AND ARE UBIQUITOUSLY EXPRESSED
CC       IN ALL EUKARYOTIC CELLS.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
CC       6.8, ITS MW IS: 46 kDa.
CC   -!- SIMILARITY: Belongs to the actin family.
DR   InterPro; IPR004001; Actin.
DR   InterPro; IPR004000; Actin_like.
DR   PROSITE; PS00406; ACTINS_1; PARTIAL.
DR   PROSITE; PS00432; ACTINS_2; PARTIAL.
DR   PROSITE; PS01132; ACTINS_ACT_LIKE; PARTIAL.
KW   Structural protein.
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  976 MW;  1424005AB2CAAEB3 CRC64;

  Query Match             26.7%;  Score 12;  DB 1;  Length 8;
  Best Local Similarity   25.0%;  Pred. No. 1.3e+05;
  Matches    1;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DINL 9
              |:::
Db          3 DVDI 6


RESULT 15
CCKN_MACEU
ID   CCKN_MACEU     STANDARD;      PRT;     8 AA.
AC   P30369;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Cholecystokinin (CCK).
GN   CCK.
OS   Macropus eugenii (Tammar wallaby), and
OS   Dasyurus viverrinus (Southeastern quoll).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX   NCBI_TaxID=9315, 9279;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=M.eugenii, and D.viverrinus;
RC   TISSUE=Brain;
RX   MEDLINE=88234141; PubMed=3375140;
RA   Fan Z.W., Eng J., Shaw G., Yalow R.S.;
RT   "Cholecystokinin octapeptide purified from brains of Australian
RT   marsupials.";
RL   Peptides 9:429-431(1988).
CC   -!- FUNCTION: THIS PEPTIDE HORMONE INDUCES GALL BLADDER CONTRACTION
CC       AND THE RELEASE OF PANCREATIC ENZYMES IN THE GUT. ITS FUNCTION
CC       IN THE BRAIN IS NOT CLEAR.
CC   -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
DR   PIR; A43001; A43001.
DR   PIR; PQ0012; PQ0012.
DR   InterPro; IPR001651; Gastrin.
DR   PROSITE; PS00259; GASTRIN; 1.
KW   Amidation; Sulfation; Hormone.
FT   MOD_RES       2      2       SULFATION.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  1064 MW;  DDCAA68378768B5A CRC64;

  Query Match             26.7%;  Score 12;  DB 1;  Length 8;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 YLSGAD 6
              |:   |
Db          2 YMGWMD 7


Search completed: January 12, 2004, 14:29:02
Job time : 6.25 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         January 12, 2004, 14:25:20 ; Search time 10.25 Seconds
                                            (without alignments)
                                            84.441 Million cell updates/sec

Title:          US-09-529-121A-4
Perfect score:  45
Sequence:       1 YLSGANINL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters:       789

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_76:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES

               %
Result       Query
  No.  Score Match Length DB ID                      Description
-------------------------------------------------------------------------
    1    16   35.6     8  2  E60588                  sperm-activating p
    2    16   35.6     8  2  H41978                  calliFMRFamide 8 -
    3    15   33.3     8  2  T13818                  cytochrome oxidase
    4    15   33.3     9  2  A61364                  isotocin - common
    5    15   33.3     9  2  PT0288                  Ig heavy chain CRD
    6    15   33.3     9  2  G41946                  T-cell receptor ga
    7    14   31.1     5  2  S62883                  seminal plasma pro
    8    14   31.1     6  2  PT0605                  T-cell receptor be
    9    14   31.1     6  2  PT0593                  T-cell receptor be
   10    14   31.1     7  2  A34818                  vicilin 72K chain
   11    14   31.1     7  2  PT0654                  T-cell receptor be
   12    14   31.1     7  2  S58797                  serine/threonine-s
   13    14   31.1     8  2  A21440                  variant surface gl
   14    14   31.1     8  2  PN0043                  phosphatidylethano
   15    14   31.1     9  2  C57444                  neuropeptide Grb-A
   16    14   31.1     9  2  JQ0914                  MHC class I histoc
   17    14   31.1     9  2  S56004                  glucan 1,3-beta-gl
   18    13   28.9     4  2  S43959                  Ig mu chain V regi
   19    13   28.9     7  2  I40504                  hypothetical prote
   20    13   28.9     7  2  I50210                  gene c-rel protein
   21    13   28.9     9  1  YFPG                    thymic factor - pi
   22    13   28.9     9  2  A60957                  thymocyte growth p
   23    13   28.9     9  2  C41170                  photosystem II pro
   24    13   28.9     9  2  PT0268                  Ig heavy chain CRD
   25    12   26.7     6  2  I51434                  H4 histone - Afric
   26    12   26.7     7  1  XEYDGD                  galactose oxidase
   27    12   26.7     8  2  PL0184                  capsid protein VP-
   28    12   26.7     8  2  PQ0701                  unidentified 6.5/3
   29    12   26.7     9  2  A44873                  caldesmon - rabbit
   30    11   24.4     4  2  S43014                  hypothetical prote
   31    11   24.4     5  1  HOROHA                  proctolin - Americ
   32    11   24.4     5  2  C23751                  spinal cord peptid
   33    11   24.4     5  2  A41225                  copper resistance
   34    11   24.4     5  2  A60411                  proctolin - Atlant
   35    11   24.4     5  2  G44817                  27.5 kda structura
   36    11   24.4     5  2  I44817                  27.5K structural p
   37    11   24.4     5  2  E44817                  27.5K structural p
   38    11   24.4     5  2  C44817                  28.5K structural p
   39    11   24.4     5  2  A44817                  28K structural pro
   40    11   24.4     6  2  B44510                  hypothetical prote
   41    11   24.4     6  2  JU0355                  lipopeptide WS1279
   42    11   24.4     6  2  PT0280                  Ig heavy chain CRD
   43    11   24.4     6  2  I49424                  cytotoxic T-lympho
   44    11   24.4     7  2  S45311                  microcin C7 - Esch
   45    11   24.4     7  2  PQ0728                  unidentified 5.0/1

                                  ALIGNMENTS

RESULT 1
E60588
sperm-activating peptide a - sea urchin (Pseudoboletia maculata)
N;Alternate names: speract homolog
C;Species: Pseudoboletia maculata
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Aug-2000
C;Accession: E60588
R;Yoshino, K.I.; Kajiura, H.; Nomura, K.; Takao, T.; Shimonishi, Y.; Kurita, M.; Yamaguc
Comp. Biochem. Physiol. B 94, 739-751, 1989
A;Title: A halogenated amino acid-containing sperm activating peptide and its related pe
otus nudus, Echinometra mathaei and Heterocentrotus mammillatus.
A;Reference number: A60527
A;Accession: E60588
A;Molecule type: protein
A;Residues: 1-8 <YOS>
C;Superfamily: unassigned animal peptides

  Query Match             35.6%;  Score 16;  DB 2;  Length 8;
  Best Local Similarity   60.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 LSGAN 6
              | | |
Db          4 LDGVN 8

RESULT 2
H41978
calliFMRFamide 8 - bluebottle fly (Calliphora vomitoria)
C;Species: Calliphora vomitoria
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
C;Accession: H41978
R;Duve, H.; Johnsen, A.H.; Sewell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, J.F.; Thorpe,
Proc. Natl. Acad. Sci. U.S.A. 89, 2326-2330, 1992
A;Title: Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH-2 neuropeptides (desi
A;Reference number: A41978; MUID:92196111; PMID:1549595
A;Accession: H41978
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-8 <DUV>
C;Keywords: amidated carboxyl end; neuropeptide
F;8/Modified site: amidated carboxyl end (Phe) #status experimental

  Query Match             35.6%;  Score 16;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GAN 6
              |||
Db          1 GAN 3
```

```
RESULT 3
T13818
cytochrome oxidase subunit I - Atlantic hagfish mitochondrion (fragment)
C;Species: mitochondrion Myxine glutinosa (Atlantic hagfish)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T13818
R;Delarbre, C.; Barriel, V.; Tillier, S.; Janvier, P.; Gachelin, G.
Mol. Biol. Evol. 14, 807-813, 1997
A;Title: The main features of the craniate mitochondrial DNA between the ND1 and the COI
A;Reference number: Z17775; MUID:97398704; PMID:9254918
A;Accession: T13818
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-8 <DEL>
A;Cross-references: EMBL:Y09527; NID:g2340019; PIDN:CAA70718.1; PID:g2340022
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion

  Query Match             33.3%;  Score 15;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLS 3
              |||
Db          2 YLS 4


RESULT 4
A61364
isotocin - common carp
C;Species: Cyprinus carpio (common carp)
C;Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 20-Jun-2000
C;Accession: A61364
R;Acher, R.; Chauvet, J.; Chauvet, M.T.; Crepy, D.
Comp. Biochem. Physiol. A 14, 245-254, 1965
A;Title: Caracterisation des hormones neurohypophysaires d'un poisson osseux d'eau douce
A;Reference number: A61364
A;Accession: A61364
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <ACH>
C;Superfamily: oxytocin-neurophysin
C;Keywords: amidated carboxyl end; neuropeptide; posterior pituitary
F;9/Modified site: amidated carboxyl end (Gly) #status experimental

  Query Match             33.3%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   42.9%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 YLSGANI 7
              |:|   |
Db          2 YISNCPI 8


RESULT 5
PT0288
Ig heavy chain CRD3 region (clone 4-106) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0288
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0288
A;Molecule type: DNA
A;Residues: 1-9 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

  Query Match             33.3%;  Score 15;  DB 2;  Length 9;
```

```
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSG 4
              | ||
Db          5 YSSG 8


RESULT 6
G41946
T-cell receptor gamma chain (2t.23) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C;Accession: G41946
R;Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A;Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ger
A;Reference number: A41946; MUID:92049316; PMID:1658619
A;Accession: G41946
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-9 <WHE>
C;Keywords: T-cell receptor

  Query Match             33.3%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSG 4
              | ||
Db          5 YSSG 8


RESULT 7
S62883
seminal plasma protein II - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C;Accession: S62883
R;Romero, A.; Varela, P.F.; Sanz, L.; Toepfer-Petersen, E.; Calvete, J.J.
FEBS Lett. 382, 15-17, 1996
A;Title: Crystallization and preliminary X-ray diffraction analysis of boar seminal plasm
A;Reference number: S62882; MUID:96196555; PMID:8612739
A;Accession: S62883
A;Molecule type: protein
A;Residues: 1-5 <ROM>
C;Complex: heterodimer; seminal plasma protein I and seminal plasma protein II
C;Keywords: glycoprotein; heterodimer; semen

  Query Match             31.1%;  Score 14;  DB 2;  Length 5;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 ANIN 8
              | ||
Db          1 ARIN 4


RESULT 8
PT0605
T-cell receptor beta chain V-D-J region (120-1L) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0605
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0605
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
```

```
C;Keywords: T-cell receptor

  Query Match             31.1%;  Score 14;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 SGA 5
              |||
Db          2 SGA 4


RESULT 9
PT0593
T-cell receptor beta chain V-D-J region (159-1F) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0593
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0593
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: day 19 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             31.1%;  Score 14;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 SGA 5
              |||
Db          4 SGA 6


RESULT 10
A34818
vicilin 72K chain - pigeon pea (fragment)
C;Species: Cajanus cajan (pigeon pea)
C;Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 30-Sep-1993
C;Accession: A34818
R;Mawal, Y.R.; Mawal, M.R.; Ranjekar, P.K.
Biochem. Biophys. Res. Commun. 166, 1446-1452, 1990
A;Title: Unusual denaturation properties of vicilin from Cajanus cajan.
A;Reference number: A34818; MUID:90165956; PMID:2306256
A;Accession: A34818
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MAW>

  Query Match             31.1%;  Score 14;  DB 2;  Length 7;
  Best Local Similarity   40.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GANIN 8
              || ::
Db          1 GARVD 5


RESULT 11
PT0654
T-cell receptor beta chain V-D-J region (121-1BK) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0654
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0654
A;Status: translation not shown
```

```
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: day 4 postnatal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             31.1%;  Score 14;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 SGA 5
              |||
Db          2 SGA 4


RESULT 12
S58797
serine/threonine-specific protein kinase c-mos - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 23-Feb-1997
C;Accession: S58797
R;Nagao, Y.
Biochim. Biophys. Acta 1245, 130-143, 1995
A;Title: Expression of c-mos protein in cultured rat spermatogenic cells and evidence tha
A;Reference number: S58797; MUID:95383384; PMID:7654761
A;Accession: S58797
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <NAG>
C;Genetics:
A;Gene: c-mos
C;Keywords: phosphotransferase; protein kinase

  Query Match             31.1%;  Score 14;  DB 2;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GANI 7
              | |:
Db          3 GGNL 6


RESULT 13
A21440
variant surface glycoprotein pSLc1 - Trypanosoma brucei (fragment)
C;Species: Trypanosoma brucei
C;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 20-Mar-1998
C;Accession: A21440
R;Parsons, M.; Nelson, R.G.; Watkins, K.P.; Agabian, N.
Cell 38, 309-316, 1984
A;Title: Trypanosome mRNAs share a common 5' spliced leader sequence.
A;Reference number: A90853; MUID:84282716; PMID:6088073
A;Accession: A21440
A;Molecule type: mRNA
A;Residues: 1-8 <PAR>
A;Cross-references: GB:K02195; NID:g162150; PID:g162151
C;Keywords: glycoprotein

  Query Match             31.1%;  Score 14;  DB 2;  Length 8;
  Best Local Similarity   33.3%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 LSGANI 7
              :||  :
Db          1 MSGKEV 6


RESULT 14
PN0043
phosphatidylethanol amine-binding protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 29-Oct-1997 #sequence_revision 29-Oct-1997 #text_change 23-Jan-1998
C;Accession: PN0043
R;Kato, H.
```

```
Kawasaki Igakkaishi 22, 245-259, 1996
A;Title: Analysis of proteins isolated by two dimensional electrophoresis of mouse neurc
A;Reference number: PN0041
A;Accession: PN0043
A;Molecule type: protein
A;Residues: 1-8 <KAT>
A;Experimental source: neuroblastoma cell
C;Comment: The molecular mass is 23,500 and the pI is 4.8. The amino-terminus is blocked
C;Keywords: brain

  Query Match            31.1%;  Score 14;  DB 2;  Length 8;
  Best Local Similarity  100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |||
Db          5 LSG 7


RESULT 15
C57444
neuropeptide Grb-AST B3 - two-spotted cricket
C;Species: Gryllus bimaculatus (two-spotted cricket)
C;Date: 26-Jan-1996 #sequence_revision 26-Jan-1996 #text_change 26-Jan-1996
C;Accession: C57444
R;Lorenz, M.W.; Kellner, R.; Hoffmann, K.H.
J. Biol. Chem. 270, 21103-21108, 1995
A;Title: A family of neuropeptides that inhibit juvenile hormone biosynthesis in the cri
A;Reference number: A57444; MUID:95403341; PMID:7673141
A;Accession: C57444
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <LOR>

  Query Match            31.1%;  Score 14;  DB 2;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |||
Db          5 LSG 7


Search completed: January 12, 2004, 14:31:53
Job time : 11.25 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          January 12, 2004, 14:21:34 ; Search time 6.25 Seconds
                                             (without alignments)
                                             67.718 Million cell updates/sec

Title:           US-09-529-121A-4
Perfect score:   45
Sequence:        1 YLSGANINL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters:      251

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_41:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                        SUMMARIES
                      %
Result               Query
  No.    Score  Match Length DB  ID                        Description
-------------------------------------------------------------------------------
     1      17   37.8      8  1  CPD1_ENTFA                P13269 enterococcu
     2      16   35.6      8  1  FAR8_CALVO                P41863 calliphora
     3      15   33.3      9  1  ISOT_CYPCA                P42993 cyprinus ca
     4      13   28.9      9  1  OXYT_RAJCL                P42994 raja clavat
     5      13   28.9      9  1  THYF_PIG                  P01255 sus scrofa
     6      12   26.7      7  1  IGAO_DACDE                P06294 dactylium d
     7      12   26.7      9  1  BS43_SERPL                P83375 serratia pl
     8      12   26.7      9  1  PGLR_DIAAB                P81179 diaprepes a
     9      11   24.4      4  1  FAR3_HIRME                P42562 hirudo medi
    10      11   24.4      5  1  PRCT_PERAM                P01373 periplaneta
    11      11   24.4      7  1  FAR2_ASCSU                P31890 ascaris suu
    12      11   24.4      7  1  GFRP_MOUSE                P99025 mus musculu
    13      11   24.4      7  1  LANC_CARUI                P36960 carnobacter
    14      11   24.4      8  1  CAD1_ENTFA                P13268 enterococcu
    15      11   24.4      8  1  LCK8_LEUMA                P19990 leucophaea
    16      11   24.4      9  1  FAR6_MACRS                P83279 macrobrachi
    17      11   24.4      9  1  MOSF_CLYJA                P19853 clypeaster
    18      11   24.4      9  1  OXYA_SCYCA                P42996 scyliorhinu
    19      11   24.4      9  1  OXYT_BUFRE                P42995 bufo regula
    20      11   24.4      9  1  TAL3_PICJA                P17441 pichia jadi
    21      10   22.2      5  1  TPIS_CANFA                P54714 canis famil
    22      10   22.2      5  1  UXA4_CHLTR                P38005 chlamydia t
    23      10   22.2      6  1  CIP2_MYTED                P13737 mytilus edu
    24      10   22.2      6  1  TMOF_SARBU                P41495 sarcophaga
    25      10   22.2      7  1  ALL2_CARMA                P81805 carcinus ma
    26      10   22.2      7  1  ALL3_CARMA                P81806 carcinus ma
    27      10   22.2      7  1  ALL4_CARMA                P81807 carcinus ma
    28      10   22.2      7  1  ALL5_CARMA                P81808 carcinus ma
    29      10   22.2      7  1  UN06_PINPS                P81675 pinus pinas
    30      10   22.2      8  1  AL12_CARMA                P81815 carcinus ma
    31      10   22.2      8  1  AL17_CARMA                P81820 carcinus ma
    32      10   22.2      8  1  ALL1_CYDPO                P82152 cydia pomon
    33      10   22.2      8  1  ALL6_CYDPO                P82157 cydia pomon
    34      10   22.2      8  1  ALL7_CARMA                P81809 carcinus ma
    35      10   22.2      8  1  ALL8_CARMA                P81811 carcinus ma
    36      10   22.2      8  1  ALL9_CARMA                P81812 carcinus ma
    37      10   22.2      8  1  B44K_PORGI                P81886 porphyromon
    38      10   22.2      8  1  GLUR_HUMAN                P02729 homo sapien
    39      10   22.2      8  1  LCK3_LEUMA                P21142 leucophaea
    40      10   22.2      8  1  LCK5_LEUMA                P19987 leucophaea
    41      10   22.2      8  1  LPMS_STAEP                P23211 staphylococ
    42      10   22.2      8  1  PLP_BRANA                 P81707 brassica na
    43      10   22.2      8  1  UF06_MOUSE                P38644 mus musculu
    44      10   22.2      8  1  UPA1_HUMAN                P30087 homo sapien
    45      10   22.2      8  1  VGLG_HSV2B                P81780 herpes simp

                                 ALIGNMENTS


RESULT 1
CPD1_ENTFA
ID   CPD1_ENTFA     STANDARD;      PRT;     8 AA.
AC   P13269;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Sex pheromone cPD1.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=85040388; PubMed=6436978;
RA   Suzuki A., Mori M., Sagakami Y., Isogai A., Fujino M., Kitada C.,
RA   Craig R.A., Clewell D.B.;
RT   "Isolation and structure of bacterial sex pheromone, cPD1.";
RL   Science 226:849-850(1984).
CC   -!- FUNCTION: CPD1 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC       BACTERIOCIN PLASMID PPD1.
KW   Pheromone.
SQ   SEQUENCE   8 AA;  913 MW;  8665B729C682C729 CRC64;

  Query Match             37.8%;  Score 17;  DB 1;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSG 4
              :|||
Db          5 FLSG 8


RESULT 2
FAR8_CALVO
ID   FAR8_CALVO     STANDARD;      PRT;     8 AA.
AC   P41863;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   CalliFMRFamide 8.
OS   Calliphora vomitoria (Blue blowfly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC   Calliphoridae; Calliphora.
OX   NCBI_TaxID=27454;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Thoracic ganglion;
RX   MEDLINE=92196111; PubMed=1549595;
RA   Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA   Rehfeld J.F., Thorpe A.;
RT   "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT   neuropeptides (designated calliFMRFamides) from the blowfly
RT   Calliphora vomitoria.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
```

```
CC       FAMILY.
DR   PIR; H41978; H41978.
KW   Neuropeptide; Amidation.
FT   MOD_RES      8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  957 MW;  72D40699CAA44DD8 CRC64;

  Query Match             35.6%;  Score 16;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GAN 6
              |||
Db          1 GAN 3


RESULT 3
ISOT_CYPCA
ID   ISOT_CYPCA     STANDARD;      PRT;     9 AA.
AC   P42993;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Isotocin.
OS   Cyprinus carpio (Common carp).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Cyprinus.
OX   NCBI_TaxID=7962;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Pituitary;
RA   Acher R., Chauvet J., Chauvet M.-T., Crepy D.;
RT   "Characterization of neurohypophyseal hormones from a fresh water bony
RT   fish, the carp (Cyprinus carpio). Comparison with hormones from sea
RT   water bony fishs.";
RL   Comp. Biochem. Physiol. 14:245-254(1965).
CC   -!- FUNCTION: ANTIDIURETIC HORMONE.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   PIR; A61364; A61364.
DR   InterPro; IPR000981; Neurhyp_horm.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Amidation.
FT   DISULFID      1      6
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  969 MW;  17FF476EB455B04B CRC64;

  Query Match             33.3%;  Score 15;  DB 1;  Length 9;
  Best Local Similarity   42.9%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 YLSGANI 7
              |:|   |
Db          2 YISNCPI 8


RESULT 4
OXYT_RAJCL
ID   OXYT_RAJCL     STANDARD;      PRT;     9 AA.
AC   P42994;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Glumitocin.
OS   Raja clavata (Thornback ray).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;
OC   Rajiformes; Rajidae; Raja.
OX   NCBI_TaxID=7781;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=66123415; PubMed=5880565;
RA   Acher R., Chauvet J., Chauvet M.-T., Crepy D.;
RT   "Phylogeny of neurophyophyseal peptides: isolation of a new hormone,
RT   glumitocin (Ser 4-Gln 8-ocytocin) present in a cartilaginous fish,
RT   the ray (Raia clavata).";
RL   Biochim. Biophys. Acta 107:393-396(1965).
CC   -!- FUNCTION: ANTIDIURETIC HORMONE.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   InterPro; IPR000981; Neurhyp_horm.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Amidation.
FT   DISULFID      1      6
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  984 MW;  17E9C76EB455B04B CRC64;

  Query Match             28.9%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLS 3
              |:|
Db          2 YIS 4


RESULT 5
THYF_PIG
ID   THYF_PIG       STANDARD;      PRT;     9 AA.
AC   P01255;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Thymic factor.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=78026571; PubMed=914862;
RA   Pleau J.-M., Dardenne M., Blouquit Y., Bach J.-F.;
RT   "Structural study of circulating thymic factor: a peptide isolated
RT   from pig serum. II. Amino acid sequence.";
RL   J. Biol. Chem. 252:8045-8047(1977).
CC   -!- MISCELLANEOUS: THE BIOLOGICAL SOURCE(S) AND PHYSIOLOGICAL
CC       ACTIVITIES OF THIS PEPTIDE HAVE NOT BEEN DETERMINED.
DR   PIR; A01523; YFPG.
KW   Pyrrolidone carboxylic acid.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
SQ   SEQUENCE   9 AA;  876 MW;  D500B87866C5B33D CRC64;

  Query Match             28.9%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GAN 6
              |:|
Db          7 GSN 9


RESULT 6
IGAO_DACDE
ID   IGAO_DACDE     STANDARD;      PRT;     7 AA.
AC   P06294;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   Galactose oxidase inhibitor.
OS   Dactylium dendroides (Cladobotryum dendroides).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypomyces.
OX   NCBI_TaxID=5132;
RN   [1]
RP   SEQUENCE.
RA   Avigad G., Markus Z.;
```

```
RT   "Identification of a peptide inhibitor of galactose oxidase from
RT   Dactylium dendroides.";
RL   Fed. Proc. 31:447-447(1972).
CC   -!- FUNCTION: BINDS ONE COPPER ION PER MOLECULE BUT DOES NOT BIND THE
CC       GALACTOSE OXIDASE APOENZYME. IT MAY INACTIVATE THE ENZYME BY
CC       BINDING TO ITS PROSTHETIC COPPER GROUP.
DR   PIR; A01341; XEYDGD.
KW   Copper; Metalloenzyme inhibitor.
SQ   SEQUENCE   7 AA;  706 MW;  75BB01A456D87DB0 CRC64;

  Query Match             26.7%;  Score 12;  DB 1;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAN 6
              :| |
Db          1 AGQN 4


RESULT 7
BS43_SERPL
ID   BS43_SERPL     STANDARD;      PRT;     9 AA.
AC   P83375;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Bacteriocin serracin P 43 kDa subunit (Fragment).
OS   Serratia plymuthica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Serratia.
OX   NCBI_TaxID=82996;
RN   [1]
RP   SEQUENCE, AND FUNCTION.
RC   STRAIN=J7;
RX   MEDLINE=22293561; PubMed=12406768;
RA   Jabrane A., Sabri A., Compere P., Jacques P., Vandenberghe I.,
RA   Van Beeumen J., Thonart P.;
RT   "Characterization of serracin P, a phage-tail-like bacteriocin, and
RT   its activity against Erwinia amylovora, the fire blight pathogen.";
RL   Appl. Environ. Microbiol. 68:5704-5710(2002).
CC   -!- FUNCTION: Major component of a prophage tail sheath (Probable).
CC   -!- FUNCTION: Antibacterial activity against Gram-negative bacterium
CC       E.amylovora.
DR   InterPro; IPR006498; Tail_tube.
DR   Pfam; PF04985; Phage_tube; 1.
KW   Antibiotic; Bacteriocin.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1095 MW;  1E66D412C871E1FB CRC64;

  Query Match             26.7%;  Score 12;  DB 1;  Length 9;
  Best Local Similarity   28.6%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 YLSGANI 7
              |  |:
Db          2 YHHGVRV 8


RESULT 8
PGLR_DIAAB
ID   PGLR_DIAAB     STANDARD;      PRT;     9 AA.
AC   P81179;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Endo-polygalacturonase (PG) (EC 3.2.1.15) (Fragment).
OS   Diaprepes abbreviatus (Sugarcane rootstalk borer weevil).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC   Phytophaga; Curculionidae; Entiminae; Entimini; Diaprepes.
OX   NCBI_TaxID=13040;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Larval gut;
RA   Doostdar H., McCollum T.G., Mayer R.T.;
RT   "Purification and characterization of an endo-polygalacturonase from
RT   the gut of West Indies sugarcane rootstalk borer weevil (Diaprepes
RT   abbreviatus L.) larvae.";
RL   Comp. Biochem. Physiol. 118B:861-867(1997).
CC   -!- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
CC       galactosiduronic linkages in pectate and other galacturonans.
CC   -!- INDUCTION: INHIBITED BY CITRUS PGIP.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
CC       9.4, ITS MW IS: 44.5 kDa.
CC   -!- SIMILARITY: WEAK, TO OTHER POLYGALACTURONASES.
KW   Hydrolase; Glycosidase; Cell wall.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1041 MW;  1F49087042DB41BB CRC64;

  Query Match             26.7%;  Score 12;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSG 4
              |: |
Db          4 YVIG 7


RESULT 9
FAR3_HIRME
ID   FAR3_HIRME     STANDARD;      PRT;     4 AA.
AC   P42562;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFamide-like neuropeptide YLRF-amide.
OS   Hirudo medicinalis (Medicinal leech).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92195954; PubMed=1686933;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Peptides 12:897-908(1991).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       4      4       AMIDATION.
SQ   SEQUENCE   4 AA;  598 MW;  69D4073B30000000 CRC64;

  Query Match             24.4%;  Score 11;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YL 2
              ||
Db          1 YL 2


RESULT 10
PRCT_PERAM
ID   PRCT_PERAM     STANDARD;      PRT;     5 AA.
AC   P01373;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   Proctolin.
OS   Periplaneta americana (American cockroach),
OS   Limulus polyphemus (Atlantic horseshoe crab), and
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattoidea;
OC   Blattidae; Periplaneta.
OX   NCBI_TaxID=6978, 6850, 6759;
```

```
RN   [1]
RP   SEQUENCE.
RC   SPECIES=P.americana;
RX   MEDLINE=76074708; PubMed=576;
RA   Starratt A.N., Brown B.E.;
RT   "Structure of the pentapeptide proctolin, a proposed neurotransmitter
RT   in insects.";
RL   Life Sci. 17:1253-1256(1975).
RN   [2]
RP   BIOLOGICAL SOURCE.
RC   SPECIES=P.americana;
RX   MEDLINE=81225865; PubMed=6113690;
RA   O'Shea M., Adams M.E.;
RT   "Pentapeptide (proctolin) associated with an identified neuron.";
RL   Science 213:567-569(1981).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=L.polyphemus;
RX   MEDLINE=90287800; PubMed=2356151;
RA   Groome J.R., Tillinghast E.K., Townley M.A., Vetrovs A.,
RA   Watson W.H. III, Hunt D.F., Griffin P.R., Alexander J.E.,
RA   Shabanowitz J.;
RT   "Identification of proctolin in the central nervous system of the
RT   horseshoe crab, Limulus polyphemus.";
RL   Peptides 11:205-211(1990).
RN   [4]
RP   SEQUENCE.
RC   SPECIES=C.maenas;
RX   MEDLINE=86232789; PubMed=2872661;
RA   Stangier J., Dircksen H., Keller R.;
RT   "Identification and immunocytochemical localization of proctolin in
RT   pericardial organs of the shore crab, Carcinus maenas.";
RL   Peptides 7:67-72(1986).
CC   -!- FUNCTION: STIMULATES CARDIAC OUTPUT AND HINDGUT MOTILITY,
CC       MODULATES VISCERAL AND SKELETAL MUSCLE IN MANY ARTHROPODS.
CC   -!- TISSUE SPECIFICITY: FOUND IN THE LATERAL WHITE NEURONS AND IN
CC       THE CRAB PERICARDIAL ORGANS.
DR   PIR; A01644; HOROHA.
DR   PIR; A60411; A60411.
KW   Neuropeptide.
SQ   SEQUENCE   5 AA;  649 MW;  71B7673B44600000 CRC64;

  Query Match             24.4%;  Score 11;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YL 2
              ||
Db          2 YL 3


RESULT 11
FAR2_ASCSU
ID   FAR2_ASCSU     STANDARD;      PRT;     7 AA.
AC   P31890;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   FMRFamide-like neuropeptide AF2.
OS   Ascaris suum (Pig roundworm) (Ascaris lumbricoides), and
OS   Panagrellus redivivus.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC   Ascarididae; Ascaris.
OX   NCBI_TaxID=6253, 6233;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=A.suum;
RX   MEDLINE=93324431; PubMed=8332542;
RA   Cowden C., Stretton A.O.W.;
RT   "AF2, an Ascaris neuropeptide: isolation, sequence, and bioactivity.";
RL   Peptides 14:423-430(1993).
RN   [2]
RP   SEQUENCE.
```

```
RC   SPECIES=P.redivivus;
RX   MEDLINE=95060998; PubMed=7970891;
RA   Maule A.G., Shaw C., Bowman J.W.;
RT   "The FMRFamide-like neuropeptide AF2 (Ascaris suum) is present in the
RT   free-living nematode, Panagrellus redivivus (Nematoda, Rhabditida).";
RL   Parasitology 109:351-356(1994).
CC   -!- FUNCTION: HAS EFFECTS ON MUSCLE TENSION.
CC   -!- TISSUE SPECIFICITY: FOUND IN THE NERVE CORDS AND A VARIETY OF
CC       GANGLIA PARTICULARLY IN THE ANTERIOR REGIONS.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       7      7       AMIDATION.
SQ   SEQUENCE   7 AA;  992 MW;  69D4073B5B11E350 CRC64;

  Query Match             24.4%;  Score 11;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YL 2
              ||
Db          4 YL 5


RESULT 12
GFRP_MOUSE
ID   GFRP_MOUSE     STANDARD;      PRT;     7 AA.
AC   P99025;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   GTP cyclohydrolase I feedback regulatory protein (P35) (Fragment).
GN   GCHFR OR GFRP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Liver;
RA   Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,
RA   Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
RA   Cowthorne M.;
RL   Submitted (AUG-1998) to the SWISS-PROT data bank.
CC   -!- FUNCTION: MEDIATES TETRAHYDROBIOPTERIN INHIBITION OF GTP
CC       CYCLOHYDROLASE I. THIS INHIBITION IS REVERSED BY L-PHENYLALANINE
CC       (BY SIMILARITY).
CC   -!- SUBUNIT: Homodimer (By similarity).
DR   SWISS-2DPAGE; P99025; MOUSE.
FT   INIT_MET      0      0
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  806 MW;  71B5B057273B4700 CRC64;

  Query Match             24.4%;  Score 11;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YL 2
              ||
Db          2 YL 3


RESULT 13
LANC_CARUI
ID   LANC_CARUI     STANDARD;      PRT;     7 AA.
AC   P36960;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Lantibiotic carnocin UI49 (Fragment).
OS   Carnobacterium sp. (strain UI49).
OC   Bacteria; Firmicutes; Lactobacillales; Carnobacteriaceae;
OC   Carnobacterium.
```

```
OX   NCBI_TaxID=35782;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92321768; PubMed=1622206;
RA   Stoffels G., Nissen-Meyer J., Gudmundsdottir A., Sletten K., Holo H.,
RA   Nes I.F.;
RT   "Purification and characterization of a new bacteriocin isolated from
RT   a Carnobacterium sp.";
RL   Appl. Environ. Microbiol. 58:1417-1422(1992).
CC   -!- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC).
CC       ACTIVE ON GRAM-POSITIVE BACTERIA.
KW   Antibiotic; Bacteriocin; Lantibiotic.
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  786 MW;  741776D05B05B810 CRC64;

  Query Match             24.4%;  Score 11;  DB 1;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GANI 7
              |: |
Db          1 GSEI 4


RESULT 14
CAD1_ENTFA
ID   CAD1_ENTFA     STANDARD;      PRT;     8 AA.
AC   P13268;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Sex pheromone CAD1.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=85051889; PubMed=6437872;
RA   Mori M., Sagakami Y., Narita M., Isogai A., Fujino M., Kitada C.,
RA   Craig R.A., Clewell D.B., Suzuki A.;
RT   "Isolation and structure of the bacterial sex pheromone, cAD1, that
RT   induces plasmid transfer in Streptococcus faecalis.";
RL   FEBS Lett. 178:97-100(1984).
CC   -!- FUNCTION: CAD1 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC       HEMOLYSIN PLASMID PAD1.
KW   Pheromone.
SQ   SEQUENCE   8 AA;  819 MW;  047DD732C735B9C7 CRC64;

  Query Match             24.4%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |:|
Db          6 LAG 8


RESULT 15
LCK8_LEUMA
ID   LCK8_LEUMA     STANDARD;      PRT;     8 AA.
AC   P19990;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Leucokinin VIII (L-VIII).
OS   Leucophaea maderae (Madeira cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
OC   Blaberidae; Leucophaea.
OX   NCBI_TaxID=6988;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Head;
RA   Holman G.M., Cook B.J., Nachman R.J.;
RT   "Isolation, primary structure and synthesis of leucokinins VII and
RT   VIII: the final members of this new family of cephalomyotropic
RT   peptides isolated from head extracts of Leucophaea maderae.";
RL   Comp. Biochem. Physiol. 88C:31-34(1987).
CC   -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC       ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC   -!- SIMILARITY: TO THE OTHER LEUCOKININS.
DR   PIR; JS0318; JS0318.
KW   Neuropeptide; Amidation.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  902 MW;  736365AB59CAADD8 CRC64;

  Query Match             24.4%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GAN 6
              ||:
Db          1 GAD 3


Search completed: January 12, 2004, 14:29:03
Job time : 7.25 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         January 12, 2004, 14:24:44 ; Search time 25.75 Seconds
                                             (without alignments)
                                             90.193 Million cell updates/sec

Title:          US-09-529-121A-4
Perfect score:  45
Sequence:       1 YLSGANINL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters:      775

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_23:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_virus:*
                13:  sp_vertebrate:*
                14:  sp_unclassified:*
                15:  sp_rvirus:*
                16:  sp_bacteriap:*
                17:  sp_archeap:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                    SUMMARIES
                 %
Result           Query
  No.   Score    Match Length DB  ID                         Description
--------------------------------------------------------------------------------
    1      15    33.3      8  3  P87225                    P87225 saccharomyc
    2      15    33.3      8 13   P82079                    P82079 limnodynast
    3      14    31.1      7  2  Q8KMS3                    Q8kms3 klebsiella
    4      14    31.1      8  2  Q9X3K1                    Q9x3k1 prochloroco
    5      14    31.1      8  4  Q8IUB8                    Q8iub8 homo sapien
    6      14    31.1      8 13   Q9PS69                    Q9ps69 gallus gall
    7      14    31.1      9  2  Q9R7E8                    Q9r7e8 escherichia
    8      14    31.1      9  2  Q8RKC6                    Q8rkc6 erwinia chr
    9      14    31.1      9  6  Q9TRS0                    Q9trs0 oryctolagus
   10      14    31.1      9  7  Q31415                    Q31415 gallus gall
   11      14    31.1      9  10  Q8VZZ3                    Q8vzz3 pinus radia
   12      14    31.1      9  10  Q9FEC0                    Q9fec0 hordeum vul
   13      14    31.1      9  11  O35953                    O35953 mus musculu
   14      14    31.1      9  12  Q89491                    Q89491 murine minu
   15      14    31.1      9  15  Q8UTD7                    Q8utd7 human immun
   16      13    28.9      7  15  Q07624                    Q07624 rous sarcom
```

```
   17      13    28.9      8  3  Q9HDS4                    Q9hds4 aspergillus
   18      13    28.9      8 13   Q91098                    Q91098 manorina me
   19      13    28.9      8 13   Q90498                    Q90498 erythrura g
   20      13    28.9      9  2  Q44377                    Q44377 aeromonas t
   21      13    28.9      9  2  Q44468                    Q44468 aeromonas v
   22      13    28.9      9  2  Q48686                    Q48686 lactococcus
   23      13    28.9      9  2  Q43928                    Q43928 aeromonas p
   24      13    28.9      9  2  Q44001                    Q44001 aeromonas e
   25      13    28.9      9  8  Q8WIJ5                    Q8wij5 bostrychia
   26      13    28.9      9  8  Q9TJ87                    Q9tj87 caloglossa
   27      13    28.9      9  8  Q8W7T9                    Q8w7t9 bostrychia
   28      13    28.9      9  8  Q9T389                    Q9t389 caloglossa
   29      13    28.9      9  8  O78337                    O78337 caloglossa
   30      13    28.9      9  8  Q9T387                    Q9t387 bostrychia
   31      13    28.9      9  8  Q9T388                    Q9t388 caloglossa
   32      13    28.9      9  8  Q9TLD0                    Q9tld0 bostrychia
   33      13    28.9      9  8  Q8W875                    Q8w875 bostrychia
   34      13    28.9      9  8  Q8ME58                    Q8me58 porphyra li
   35      13    28.9      9  8  Q9MVU1                    Q9mvu1 caloglossa
   36      13    28.9      9  8  Q8W7U0                    Q8w7u0 bostrychia
   37      13    28.9      9  8  Q9TJ85                    Q9tj85 caloglossa
   38      13    28.9      9  8  Q95BY3                    Q95by3 caloglossa
   39      13    28.9      9  8  Q9MVU2                    Q9mvu2 caloglossa
   40      13    28.9      9  8  Q8MDU2                    Q8mdu2 hypnea sp.
   41      13    28.9      9  8  Q8ME56                    Q8me56 porphyra mi
   42      13    28.9      9  8  Q8HSL5                    Q8hsl5 spyridia fi
   43      13    28.9      9  8  Q8HSL0                    Q8hsl0 murrayella
   44      13    28.9      9  8  Q8HRK0                    Q8hrk0 gracilaria
   45      13    28.9      9  8  Q8HRJ8                    Q8hrj8 ptilophora

                                    ALIGNMENTS


RESULT 1
P87225
ID   P87225      PRELIMINARY;      PRT;     8 AA.
AC   P87225;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   GIN11 protein (Fragment).
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wedler H., Wedler E., Scharfe M., Wambutt R.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   MIPS;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z73169; CAA97518.2; -.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  1019 MW;  4E21A9C449D5B73B CRC64;

  Query Match             33.3%;  Score 15;  DB 3;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLS 3
              |||
Db          1 YLS 3


RESULT 2
P82079
ID   P82079      PRELIMINARY;      PRT;     8 AA.
AC   P82079;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel.. 13, Last sequence update)
```

```
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DYNASTIN 1.
OS   Limnodynastes interioris (Giant banjo frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Limnodynastinae; Limnodynastes.
OX   NCBI_TaxID=30362;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=TIBIAL GLAND;
RA   Raftery M.J., Bradford A.M., Bowie J.H., Wallace J.C., Tyler M.J.;
RT   "Peptides from Australian frogs. The structure of the dynastins from
RT   the banjo frogs Limnodynastes interioras, Limnodynastes dumerilii and
RT   Limnodynastes terraereginae.";
RL   Aust. J. Chem. 46:833-842(1993).
CC   -!- MASS SPECTROMETRY: MW=729; METHOD=FAB.
KW   Amphibian skin.
SQ   SEQUENCE   8 AA;  729 MW;  7C28772865B72728 CRC64;

  Query Match             33.3%;  Score 15;  DB 13;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 LSGANI 7
              |||  :
Db          3 LSGLGL 8


RESULT 3
Q8KMS3
ID   Q8KMS3      PRELIMINARY;      PRT;     7 AA.
AC   Q8KMS3;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Putative MerR2 protein.
GN   MERR2.
OS   Klebsiella sp. LS13-39.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Klebsiella.
OX   NCBI_TaxID=143776;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LS13-39;
RX   MEDLINE=21604134; PubMed=11763242;
RA   Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,
RA   Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,
RA   Yurieva O.V., Nikiforov V.G.;
RT   "Mercury resistance transposons of Gram-negative environmental
RT   bacteria and their classification.";
RL   Res. Microbiol. 152:811-822(2001).
DR   EMBL; AJ302776; CAC82975.1; -.
SQ   SEQUENCE   7 AA;  608 MW;  6DC1B5BDD87DD6F0 CRC64;

  Query Match             31.1%;  Score 14;  DB 2;  Length 7;
  Best Local Similarity   40.0%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSGAN 6
              ::||:
Db          1 MAGAS 5


RESULT 4
Q9X3K1
ID   Q9X3K1      PRELIMINARY;      PRT;     8 AA.
AC   Q9X3K1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   Cytochrome b (Fragment).
GN   PETB.
OS   Prochlorococcus sp.
OC   Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC   Prochlorococcus.
OX   NCBI_TaxID=1220;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Urbach E., Chisholm S.W.;
RT   "Genetic diversity in Prochlorococcus populations flow cytometrically
RT   sorted from the Sargasso Sea and Gulf Stream.";
RL   Limnol. Oceanog. 43:1615-1630(1998).
DR   EMBL; AF070193; AAD23233.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  799 MW;  10376865B72866D3 CRC64;

  Query Match             31.1%;  Score 14;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |||
Db          4 LSG 6


RESULT 5
Q8IUB8
ID   Q8IUB8      PRELIMINARY;      PRT;     8 AA.
AC   Q8IUB8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CD95 antigen (Fragment).
GN   CD95.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22404279; PubMed=12516573;
RA   Kurth J., Perniok A., Schmitz R., Iking-Konert C., Chiorazzi N.,
RA   Thompson K.M., Winkler T., Rajewsky K., Kueppers R.;
RT   "Lack of deleterious somatic mutations in the CD95 gene of
RT   plasmablasts from systemic lupus erythematosus patients and
RT   autoantibody-producing cell lines.";
RL   Eur. J. Immunol. 32:3785-3792(2002).
DR   EMBL; AJ509178; CAD48928.1; -.
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  846 MW;  34B724405DC2D1AB CRC64;

  Query Match             31.1%;  Score 14;  DB 4;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 INL 9
              |||
Db          5 INL 7


RESULT 6
Q9PS69
ID   Q9PS69      PRELIMINARY;      PRT;     8 AA.
AC   Q9PS69;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   LOW density lipoprotein receptor-related protein (Fragment).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE.
```

```
RX   MEDLINE=92011685; PubMed=1918027;
RA   Stifani S., Barber D.L., Aebersold R., Steyrer E., Shen X., Nimpf J.,
RA   Schneider W.J.;
RT   "The laying hen expresses two different low density lipoprotein
RT   receptor-related proteins.";
RL   J. Biol. Chem. 266:19079-19087(1991).
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  846 MW;  C007272DD865BAAA CRC64;

  Query Match             31.1%;  Score 14;  DB 13;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 SGA 5
              |||
Db          3 SGA 5


RESULT 7
Q9R7E8
ID   Q9R7E8      PRELIMINARY;      PRT;     9 AA.
AC   Q9R7E8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   KpsD (Fragment).
GN   KPSD.
OS   Escherichia coli.
OG   Plasmid pCR3.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95180691; PubMed=7875563;
RA   Rosenow C., Roberts I.S., Jann K.;
RT   "Isolation from recombinant Escherichia coli and characterization of
RT   CMP-Kdo synthetase, involved in the expression of the capsular K5
RT   polysaccharide (K-CKS).";
RL   FEMS Microbiol. Lett. 125:159-164(1995).
DR   EMBL; S76943; AAB33585.1; -.
KW   Plasmid.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  899 MW;  3EBBB72042C33DD8 CRC64;

  Query Match             31.1%;  Score 14;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 GANINL 9
              ||:  |
Db          2 GAKVIL 7


RESULT 8
Q8RKC6
ID   Q8RKC6      PRELIMINARY;      PRT;     9 AA.
AC   Q8RKC6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   ExpR protein (Fragment).
GN   EXPR.
OS   Erwinia chrysanthemi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Pectobacterium.
OX   NCBI_TaxID=556;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3937;
RA   Reverchon S.;
RT   "Identification of a lysA-like gene required for virulence factors
RT   synthesis in Erwinia chrysanthemi.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ438189; CAD27339.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  999 MW;  9A8BC455B9D5B045 CRC64;

  Query Match             31.1%;  Score 14;  DB 2;  Length 9;
  Best Local Similarity   28.6%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 LSGANIN 8
              :| :|::
Db          3 ISFSNVD 9


RESULT 9
Q9TRS0
ID   Q9TRS0      PRELIMINARY;      PRT;     9 AA.
AC   Q9TRS0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Calcyclin-associated protein, CAP50=CA2+/phospholipid-binding protein
DE   L-7 fragment (Fragment).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92250478; PubMed=1533622;
RA   Tokumitsu H., Mizutani A., Minami H., Kobayashi R., Hidaka H.;
RT   "A calcyclin-associated protein is a newly identified member of the
RT   Ca2+/phospholipid-binding proteins, annexin family.";
RL   J. Biol. Chem. 267:8919-8924(1992).
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1010 MW;  64E419C44865B72B CRC64;

  Query Match             31.1%;  Score 14;  DB 6;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |||
Db          3 LSG 5


RESULT 10
Q31415
ID   Q31415      PRELIMINARY;      PRT;     9 AA.
AC   Q31415;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC class I antigen (Fragment).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kanki T., Kuwasawa N., Sekiya Y., Ichikawa Y.;
RT   "Responsive expression of a MHC class I epitope and genes following
RT   Marek's disease virus infection.";
RL   Submitted (JAN-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; D90399; BAA14395.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  859 MW;  8A55A76455B861B5 CRC64;

  Query Match             31.1%;  Score 14;  DB 7;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+05;
```

```
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 SGAN 6
              :|:|
Db          3 TGSN 6


RESULT 11
Q8VZZ3
ID   Q8VZZ3      PRELIMINARY;      PRT;     9 AA.
AC   Q8VZZ3;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Caffeic acid O-methyltransferase (Fragment).
GN   AEOMT.
OS   Pinus radiata (Monterey pine).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX   NCBI_TaxID=3347;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Moyle R.L., Wagner A., Walter C.;
RT   "Isolation and characterization of an AEOMT promoter fragment from
RT   Pinus radiata.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY053389; AAL17620.1; -.
KW   Transferase; Methyltransferase.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  952 MW;  86EE2874469455BA CRC64;

  Query Match             31.1%;  Score 14;  DB 10;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ANIN 8
              :|:|
Db          3 SNMN 6


RESULT 12
Q9FEC0
ID   Q9FEC0      PRELIMINARY;      PRT;     9 AA.
AC   Q9FEC0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Mla1uORF (MLA13uORF 2de).
GN   MLA1.
OS   Hordeum vulgare (Barley).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Hordeum.
OX   NCBI_TaxID=4513;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Algerian;
RA   Zhou F., Kurth J., Wei F., Elliott C., Vale G., Yahiaoui N.,
RA   Keller B., Somerville S., Wise R., Schulze-lefert P.;
RT   "Cell-autonomous Expression of Barley Mla1 Confers Race-specific
RT   Resistance to the Powdery Mildew Fungus via a Rar1 Independent
RT   Signaling Pathway.";
RL   Plant Cell 0:0-0(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Halterman D.A., Wei F., Wise R.P.;
RT   "Powdery mildew-induced Mla mRNAs are alternatively spliced and
RT   contain multiple upstream open reading frames.";
RL   Plant Physiol. 0:0-0(2003).
DR   EMBL; AY009939; AAG37357.1; -.
DR   EMBL; AY009938; AAG37355.1; -.
DR   EMBL; AF523682; AAO16013.1; -.
DR   EMBL; AF523683; AAO16016.1; -.
```

```
SQ   SEQUENCE   9 AA;  1163 MW;  473E2440573B5337 CRC64;

  Query Match             31.1%;  Score 14;  DB 10;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 INL 9
              |||
Db          6 INL 8


RESULT 13
O35953
ID   O35953      PRELIMINARY;      PRT;     9 AA.
AC   O35953;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Truncated voltage-gated sodium channel alpha subunit (Fragment).
GN   SCN8A.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RIII;
RX   MEDLINE=97442476; PubMed=9295353;
RA   Plummer N.W., McBurney M.W., Meisler M.H.;
RT   "Alternative splicing of the sodium channel SCN8A predicts a truncated
RT   two-domain protein in fetal brain and non-neuronal cells.";
RL   J. Biol. Chem. 272:24008-24015(1997).
DR   EMBL; U97672; AAB80914.1; -.
DR   MGD; MGI:103169; Scn8a.
KW   Ionic channel.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  898 MW;  22D92865B735B737 CRC64;

  Query Match             31.1%;  Score 14;  DB 11;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |||
Db          5 LSG 7


RESULT 14
Q89491
ID   Q89491      PRELIMINARY;      PRT;     9 AA.
AC   Q89491;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   Hypothetical 1.1 kDa protein.
OS   Murine minute virus (Murine parvovirus).
OC   Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX   NCBI_TaxID=10794;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LYMPHOTROPIC VARIANT;
RX   MEDLINE=86115415; PubMed=3502703;
RA   Astell C.R., Gardiner E.M., Tattersall P.;
RT   "DNA sequence of the lymphotropic variant of minute virus of mice,
RT   MVM(i), and comparison with the DNA sequence of the fibrotropic
RT   prototype strain.";
RL   J. Virol. 570:656-669(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MVM(P);
RX   MEDLINE=83143341; PubMed=6298737;
RA   Astell C.R., Thomson M., Merchlinsky M., Ward D.C.;
RT   "The complete DNA sequence of minute virus of mice, an autonomous
```

```
RT   parvovirus.";
RL   Nucleic Acids Res. 11:999-1018(1983).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MVM(P);
RX   MEDLINE=86115415; PubMed=3502703;
RA   Astell C.R., Gardiner E.M., Tattersall P.;
RT   "DNA sequence of the lymphotropic variant of minute virus of mice,
RT   MVM(i), and comparison with the DNA sequence of the fibrotropic
RT   prototype strain.";
RL   J. Virol. 57:656-669(1986).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MVM(P);
RX   MEDLINE=87061199; PubMed=3783817;
RA   Morgan W.R., Ward D.C.;
RT   "Three splicing patterns are used to excise the small intron common to
RT   all minute virus of mice RNAs.";
RL   J. Virol. 60:1170-1174(1986).
DR   EMBL; M12032; AAA69570.1; -.
DR   EMBL; J02275; AAA67112.1; -.
DR   EMBL; V01115; CAA24311.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   9 AA;  1061 MW;  C3FD405863637862 CRC64;

  Query Match             31.1%;  Score 14;  DB 12;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GANI 7
              | |:
Db          6 GINV 9


RESULT 15
Q8UTD7
ID   Q8UTD7      PRELIMINARY;      PRT;     9 AA.
AC   Q8UTD7;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Vpu protein.
GN   VPU.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=00BW1471.27;
RA   Novitsky V.A., Smith U.R., Gilbert P., McLane M.F., Chigwedere P.,
RA   Williamson C., Ndung'u T., Klein I., Chang S.-Y., Peter T., Thior I.,
RA   Foley B.T., Gaolekwe S., Rybak N., Gaseitsiwe S., Vannberg F.,
RA   Marlink R., Lee T.-H., Essex M.;
RT   "HIV-1 subtype C molecular phylogeny: consensus sequence for an AIDS
RT   vaccine design.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF443091; AAL34712.1; -.
SQ   SEQUENCE   9 AA;  1102 MW;  188BD40B17272440 CRC64;

  Query Match             31.1%;  Score 14;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 INL 9
              |||
Db          2 INL 4


Search completed: January 12, 2004, 14:31:01
Job time : 27.75 secs
```

```
                         GenCore version 5.1.6
                Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         January 12, 2004, 14:25:20 ; Search time 10.25 Seconds
                                             (without alignments)
                                             84.441 Million cell updates/sec

Title:          US-09-529-121A-5
Perfect score:  48
Sequence:       1 YLSGACLNL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters:      789

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_76:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES

                 %
Result         Query
  No.   Score  Match Length DB  ID                  Description
  -------------------------------------------------------------------
    1      17   35.4      6  2  C22565              R-phycoerythrin be
    2      17   35.4      6  2  PT0652              T-cell receptor be
    3      15   31.2      5  2  F22565              R-phycoerythrin ga
    4      15   31.2      7  2  A12016              formylglycinamide
    5      15   31.2      8  2  C61512              variant surface gl
    6      15   31.2      8  2  T13818              cytochrome oxidase
    7      15   31.2      9  2  S19329              sperm-activating p
    8      15   31.2      9  2  PT0288              Ig heavy chain CRD
    9      15   31.2      9  2  G41946              T-cell receptor ga
   10      14   29.2      5  2  B45525              actin I - malaria
   11      14   29.2      6  2  PT0605              T-cell receptor be
   12      14   29.2      6  2  PT0593              T-cell receptor be
   13      14   29.2      7  2  PT0654              T-cell receptor be
   14      14   29.2      8  2  PL0184              capsid protein VP-
   15      14   29.2      8  2  PN0043              phosphatidylethano
   16      14   29.2      9  2  C57444              neuropeptide Grb-A
   17      13   27.1      4  2  S43959              Ig mu chain V regi
   18      13   27.1      6  2  I37263              Y protein - human
   19      13   27.1      6  2  I49421              laminin B1 - weste
   20      13   27.1      7  2  S71867              glutathione transf
   21      13   27.1      7  2  S38516              mabinlin II chain
   22      13   27.1      7  2  A34026              acetylcholinestera
   23      13   27.1      7  2  I50210              gene c-rel protein
   24      13   27.1      8  2  A37521              R-phycoerythrin ga
   25      13   27.1      9  2  A61364              isotocin - common
   26      13   27.1      9  2  C41170              photosystem II pro
   27      13   27.1      9  2  A29477              diuretic neuropept
   28      13   27.1      9  4  I73804              hypothetical E2 pr
   29      12   25.0      6  2  I51434              H4 histone - Afric
```

```
   30      12   25.0      7  2  PN0649              pullulanase (EC 3.
   31      12   25.0      8  2  PH1407              Ig heavy chain V r
   32      12   25.0      8  2  PQ0701              unidentified 6.5/3
   33      12   25.0      8  2  A21440              variant surface gl
   34      12   25.0      8  2  I57018              gene Cftr protein
   35      12   25.0      9  2  A44873              caldesmon - rabbit
   36      11   22.9      5  1  HOROHA              proctolin - Americ
   37      11   22.9      5  2  C23751              spinal cord peptid
   38      11   22.9      5  2  A41225              copper resistance
   39      11   22.9      5  2  A60411              proctolin - Atlant
   40      11   22.9      6  2  B44510              hypothetical prote
   41      11   22.9      6  2  PT0280              Ig heavy chain CRD
   42      11   22.9      7  2  B34818              vicilin 57K chain
   43      11   22.9      7  2  S78024              ribosomal protein
   44      11   22.9      8  2  S59622              metallothionein is
   45      11   22.9      8  2  A61467              penalbumin - Adeli

                              ALIGNMENTS


RESULT 1
C22565
R-phycoerythrin beta-1 chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C;Accession: C22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601; PMID:3886644
A;Accession: C22565
A;Molecule type: protein
A;Residues: 1-6 <KLO>

  Query Match             35.4%;  Score 17;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ACL 7
              |||
Db          3 ACL 5


RESULT 2
PT0652
T-cell receptor beta chain V-D-J region (121-1E) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0652
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0652
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: day 4 postnatal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             35.4%;  Score 17;  DB 2;  Length 6;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAC 6
              || |
Db          2 SGDC 5


RESULT 3
F22565
R-phycoerythrin gamma-A chain - red alga (Gastroclonium coulteri) (fragment)
```

```
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C;Accession: F22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601; PMID:3886644
A;Accession: F22565
A;Molecule type: protein
A;Residues: 1-5 <KLO>

  Query Match             31.2%;  Score 15;  DB 2;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GAC 6
              | |
Db          1 GTC 3


RESULT 4
A12016
formylglycinamide ribonucleotide amidotransferase (EC 2.-.-.-) - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 13-Mar-1997
C;Accession: A12016; B12016
R;Ohnoki, S.; Hong, B.S.; Buchanan, J.M.
Fed. Proc. 35, 1549, 1976
A;Title: Amino acid sequence at glutamine active site for FGAR-amidotransferase.
A;Reference number: A91459
A;Accession: A12016
A;Molecule type: protein
A;Residues: 1-7 <OHN>
A;Experimental source: liver, peptide 1
A;Accession: B12016
A;Molecule type: protein
A;Residues: 1-5 <OH2>
A;Experimental source: liver, peptide 2
C;Keywords: transferase

  Query Match             31.2%;  Score 15;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GAC 6
              | |
Db          1 GVC 3


RESULT 5
C61512
variant surface glycoprotein MITat 1.4 - Trypanosoma brucei (fragment)
C;Species: Trypanosoma brucei
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 07-May-1999
C;Accession: C61512
R;Holder, A.A.; Cross, G.A.M.
Mol. Biochem. Parasitol. 2, 135-150, 1981
A;Title: Glycopeptides from variant surface glycoproteins of Trypanosoma brucei. C-termi
A;Reference number: A61512; MUID:81172836; PMID:6163983
A;Accession: C61512
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-8 <HOL>
C;Keywords: glycoprotein

  Query Match             31.2%;  Score 15;  DB 2;  Length 8;
  Best Local Similarity   33.3%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 SGACLN 8
              : || :
Db          3 NNACKB 8
```

```
RESULT 6
T13818
cytochrome oxidase subunit I - Atlantic hagfish mitochondrion (fragment)
C;Species: mitochondrion Myxine glutinosa (Atlantic hagfish)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T13818
R;Delarbre, C.; Barriel, V.; Tillier, S.; Janvier, P.; Gachelin, G.
Mol. Biol. Evol. 14, 807-813, 1997
A;Title: The main features of the craniate mitochondrial DNA between the ND1 and the COI
A;Reference number: Z17775; MUID:97398704; PMID:9254918
A;Accession: T13818
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-8 <DEL>
A;Cross-references: EMBL:Y09527; NID:g2340019; PIDN:CAA70718.1; PID:g2340022
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion

  Query Match             31.2%;  Score 15;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLS 3
              |||
Db          2 YLS 4


RESULT 7
S19329
sperm-activating peptide SAP - sea urchin (Stomopneustes variolus)
C;Species: Stomopneustes variolus
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 18-Aug-2000
C;Accession: S19329
R;Yoshino, K.; Takao, T.; Shimonishi, Y.; Suzuki, N.
FEBS Lett. 294, 179-182, 1991
A;Title: Determination of the amino acid sequence of an intramolecular disulfide linkage-
A;Reference number: S19329; MUID:92097763; PMID:1756858
A;Accession: S19329
A;Molecule type: protein
A;Residues: 1-9 <YOS>
C;Superfamily: unassigned animal peptides
F;3-8/Disulfide bonds: #status predicted

  Query Match             31.2%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GACL 7
              | |:
Db          6 GKCV 9


RESULT 8
PT0288
Ig heavy chain CRD3 region (clone 4-106) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0288
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and jc
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0288
A;Molecule type: DNA
A;Residues: 1-9 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

  Query Match             31.2%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
```

```
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSG 4
              | ||
Db          5 YSSG 8


RESULT 9
G41946
T-cell receptor gamma chain (2t.23) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C;Accession: G41946
R;Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A;Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
A;Reference number: A41946; MUID:92049316; PMID:1658619
A;Accession: G41946
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-9 <WHE>
C;Keywords: T-cell receptor

  Query Match             31.2%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSG 4
              | ||
Db          5 YSSG 8


RESULT 10
B45525
actin I - malaria parasite (Plasmodium falciparum) (fragments)
C;Species: Plasmodium falciparum
C;Date: 03-Jun-1993 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C;Accession: B45525
R;Wesseling, J.G.; Snijders, P.J.F.; van Someren, P.; Jansen, J.; Smits, M.A.; Schoenmak
Mol. Biochem. Parasitol. 35, 167-176, 1989
A;Title: Stage-specific expression and genomic organization of the actin genes of the ma
A;Reference number: A45525; MUID:89364996; PMID:2671721
A;Accession: B45525
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-5 <WES>
A;Cross-references: GB:J03988
A;Note: the authors translated the codon GAA for residue 3 as Gly
C;Comment: The actin I gene contains no introns.

  Query Match             29.2%;  Score 14;  DB 2;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GAC 6
              | |
Db          2 GEC 4


RESULT 11
PT0605
T-cell receptor beta chain V-D-J region (120-1L) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0605
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0605
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             29.2%;  Score 14;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 SGA 5
              |||
Db          2 SGA 4


RESULT 12
PT0593
T-cell receptor beta chain V-D-J region (159-1F) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0593
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0593
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: day 19 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             29.2%;  Score 14;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 SGA 5
              |||
Db          4 SGA 6


RESULT 13
PT0654
T-cell receptor beta chain V-D-J region (121-1BK) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0654
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0654
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: day 4 postnatal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             29.2%;  Score 14;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 SGA 5
              |||
Db          2 SGA 4


RESULT 14
PL0184
capsid protein VP-1 - murine poliovirus (fragment)
C;Species: murine poliovirus, Theiler's encephalomyelitis virus
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Feb-1995
C;Accession: PL0184
R;Zurbriggen, A.; Hogle, J.M.; Fujinami, R.S.
J. Exp. Med. 170, 2037-2049, 1989
A;Title: Alteration of amino acid 101 within capsid protein VP-1 changes the pathogenicit
```

```
A;Reference number: PL0184; MUID:90063468; PMID:2479706
A;Accession: PL0184
A;Molecule type: genomic RNA
A;Residues: 1-8 <ZUR>
C;Keywords: capsid protein

  Query Match             29.2%;  Score 14;  DB 2;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 SGACLN 8
              ||   |
Db          1 SGGITN 6


RESULT 15
PN0043
phosphatidylethanol amine-binding protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 29-Oct-1997 #sequence_revision 29-Oct-1997 #text_change 23-Jan-1998
C;Accession: PN0043
R;Kato, H.
Kawasaki Igakkaishi 22, 245-259, 1996
A;Title: Analysis of proteins isolated by two dimensional electrophoresis of mouse neuro
A;Reference number: PN0041
A;Accession: PN0043
A;Molecule type: protein
A;Residues: 1-8 <KAT>
A;Experimental source: neuroblastoma cell
C;Comment: The molecular mass is 23,500 and the pI is 4.8. The amino-terminus is blocked
C;Keywords: brain

  Query Match             29.2%;  Score 14;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |||
Db          5 LSG 7


Search completed: January 12, 2004, 14:31:53
Job time : 10.25 secs
```

```
                    GenCore version 5.1.6
             Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          January 12, 2004, 14:21:34 ; Search time 6.25 Seconds
                                              (without alignments)
                                              67.718 Million cell updates/sec

Title:           US-09-529-121A-5
Perfect score:   48
Sequence:        1 YLSGACLNL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters:      251

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_41:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES

                    %
Result           Query
   No.   Score   Match Length DB ID                       Description
-------------------------------------------------------------------------------
     1      17    35.4      8  1 CPD1_ENTFA               P13269 enterococcu
     2      15    31.2      9  1 SAP_STOVA                P24047 stomopneute
     3      13    27.1      9  1 COW_CONVE                P83047 conus ventr
     4      13    27.1      9  1 DNF1_LOCMI               P16339 locusta mig
     5      13    27.1      9  1 ISOT_CYPCA               P42993 cyprinus ca
     6      13    27.1      9  1 OXYT_RAJCL               P42994 raja clavat
     7      12    25.0      9  1 PGLR_DIAAB               P81179 diaprepes a
     8      11    22.9      4  1 FAR3_HIRME               P42562 hirudo medi
     9      11    22.9      5  1 PRCT_PERAM               P01373 periplaneta
    10      11    22.9      7  1 FAR2_ASCSU               P31890 ascaris suu
    11      11    22.9      7  1 GFRP_MOUSE               P99025 mus musculu
    12      11    22.9      8  1 CAD1_ENTFA               P13268 enterococcu
    13      11    22.9      8  1 COW2_CONPU               P58785 conus purpu
    14      11    22.9      9  1 CONO_CONST               P05487 conus stria
    15      11    22.9      9  1 MOSF_CLYJA               P19853 clypeaster
    16      11    22.9      9  1 OXYA_SQUAC               P42999 squalus aca
    17      10    20.8      5  1 UXA4_CHLTR               P38005 chlamydia t
    18      10    20.8      6  1 CIP2_MYTED               P13737 mytilus edu
    19      10    20.8      6  1 TMOF_SARBU               P41495 sarcophaga
    20      10    20.8      7  1 ALL2_CARMA               P81805 carcinus ma
    21      10    20.8      7  1 ALL3_CARMA               P81806 carcinus ma
    22      10    20.8      7  1 ALL4_CARMA               P81807 carcinus ma
    23      10    20.8      7  1 ALL5_CARMA               P81808 carcinus ma
    24      10    20.8      7  1 UN06_PINPS               P81675 pinus pinas
    25      10    20.8      8  1 AKH_MELML                P25423 melolontha
    26      10    20.8      8  1 AL12_CARMA               P81815 carcinus ma
    27      10    20.8      8  1 AL17_CARMA               P81820 carcinus ma
    28      10    20.8      8  1 ALL7_CARMA               P81809 carcinus ma
    29      10    20.8      8  1 ALL8_CARMA               P81811 carcinus ma
    30      10    20.8      8  1 ALL9_CARMA               P81812 carcinus ma
    31      10    20.8      8  1 FAR8_CALVO               P41863 calliphora
    32      10    20.8      8  1 GLUR_HUMAN               P02729 homo sapien
    33      10    20.8      8  1 HTF_TENMO                P25419 tenebrio mo
    34      10    20.8      8  1 LCK5_LEUMA               P19987 leucophaea
    35      10    20.8      8  1 LCK8_LEUMA               P19990 leucophaea
    36      10    20.8      8  1 RPCH_PANBO               P08939 pandalus bo
    37      10    20.8      8  1 UF06_MOUSE               P38644 mus musculu
    38      10    20.8      8  1 VGLG_HSV2B               P81780 herpes simp
    39      10    20.8      8  1 WP1_PERAT                P83195 perkinsus a
    40      10    20.8      9  1 AL10_CARMA               P81813 carcinus ma
    41      10    20.8      9  1 AL11_CARMA               P81814 carcinus ma
    42      10    20.8      9  1 DSIP_RABIT               P01158 oryctolagus
    43      10    20.8      9  1 FAR5_ASCSU               P43170 ascaris suu
    44      10    20.8      9  1 FAR6_CALVO               P41861 calliphora
    45      10    20.8      9  1 FIBB_MACFU               P19345 macaca fusc

                              ALIGNMENTS

RESULT 1
CPD1_ENTFA
ID   CPD1_ENTFA     STANDARD;      PRT;     8 AA.
AC   P13269;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Sex pheromone cPD1.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=85040388; PubMed=6436978;
RA   Suzuki A., Mori M., Sagakami Y., Isogai A., Fujino M., Kitada C.,
RA   Craig R.A., Clewell D.B.;
RT   "Isolation and structure of bacterial sex pheromone, cPD1.";
RL   Science 226:849-850(1984).
CC   -!- FUNCTION: CPD1 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC       BACTERIOCIN PLASMID PPD1.
KW   Pheromone.
SQ   SEQUENCE   8 AA;  913 MW;  8665B729C682C729 CRC64;

  Query Match             35.4%;  Score 17;  DB 1;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSG 4
              :|||
Db          5 FLSG 8


RESULT 2
SAP_STOVA
ID   SAP_STOVA      STANDARD;      PRT;     9 AA.
AC   P24047;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-MAR-1992 (Rel. 21, Last annotation update)
DE   Sperm-activating peptide (SAP).
OS   Stomopneutes variolaris (Sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Diadematacea; Phymosomatoida; Stomechinidae;
OC   Stomopneustes.
OX   NCBI_TaxID=7663;
RN   [1]
RP   SEQUENCE, AND DISULFIDE BOND.
RC   TISSUE=Egg jelly;
RX   MEDLINE=92097763; PubMed=1756858;
RA   Yoshimo K.-I., Takao T., Shimonishi Y., Suzuki N.;
RT   "Determination of the amino acid sequence of an intramolecular
RT   disulfide linkage-containing sperm-activating peptide by tandem mass
RT   spectrometry.";
RL   FEBS Lett. 294:179-182(1991).
CC   -!- FUNCTION: CAUSE STIMULATION OF SPERM RESPIRATION AND MOTILITY
CC       THROUGH INTRACELLULAR ALKALINIZATION, TRANSIENT ELEVATIONS OF
```

```
CC       CAMP, CGMP AND CLACIUM LEVELS IN SPERM CELLS, AND TRANSIENT
CC       ACTIVATION AND SUBSEQUENT INACTIVATION OF THE MEMBRANE FORM OF
CC       GUANYLATE CYCLASE.
FT   DISULFID      3      8
SQ   SEQUENCE   9 AA;  1010 MW;  C469B3387B076EB9 CRC64;

  Query Match             31.2%;  Score 15;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GACL 7
              | |:
Db          6 GKCV 9


RESULT 3
COW_CONVE
ID   COW_CONVE      STANDARD;      PRT;     9 AA.
AC   P83047;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Contryphan-Vn.
OS   Conus ventricosus (Mediterranean cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC   Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
OX   NCBI_TaxID=117992;
RN   [1]
RP   SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC   TISSUE=Venom;
RX   MEDLINE=21547785; PubMed=11688995;
RA   Massilia G.R., Schinina M.E., Ascenzi P., Polticelli F.;
RT   "Contryphan-Vn: a novel peptide from the venom of the Mediterranean
RT   snail Conus ventricosus.";
RL   Biochem. Biophys. Res. Commun. 288:908-913(2001).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC   -!- MASS SPECTROMETRY: MW=1088.6; METHOD=MALDI.
CC   -!- SIMILARITY: BELONGS TO THE CONTRYPHAN FAMILY.
KW   Toxin; Amidation; D-amino acid.
FT   DISULFID      3      9
FT   MOD_RES       5      5       D-TRYPTOPHAN.
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1091 MW;  8D38676323676EBA CRC64;

  Query Match             27.1%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GAC 6
              | |
Db          1 GDC 3


RESULT 4
DNF1_LOCMI
ID   DNF1_LOCMI     STANDARD;      PRT;     9 AA.
AC   P16339;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Locupressin (Diuretic neuropeptide F1/F2).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC   Acridoidea; Acrididae; Oedipodinae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Suboesophageal ganglion, and Thoracic ganglion;
RX   MEDLINE=88077077; PubMed=3689410;
RA   Proux J.P., Miller C.A., Li J.P., Carney R.L., Girardie A.,
```

```
RA   Delaage M., Schooley D.A.;
RT   "Identification of an arginine vasopressin-like diuretic hormone from
RT   Locusta migratoria.";
RL   Biochem. Biophys. Res. Commun. 149:180-186(1987).
CC   -!- FUNCTION: DIURETIC HORMONE.
CC   -!- SUBUNIT: F2 IS AN ANTIPARALLEL DISULFIDE LINKED DIMER OF F1.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   PIR; A29477; A29477.
DR   InterPro; IPR000981; Neurhyp_horm.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Neuropeptide; Amidation.
FT   DISULFID      1      6       IN F1.
FT   DISULFID      1      1       INTERCHAIN (WITH C-6') (IN F2).
FT   DISULFID      6      6       INTERCHAIN (WITH C-1') (IN F2).
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  976 MW;  56EB176EB451A057 CRC64;

  Query Match             27.1%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 CL 7
              ||
Db          1 CL 2


RESULT 5
ISOT_CYPCA
ID   ISOT_CYPCA     STANDARD;      PRT;     9 AA.
AC   P42993;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Isotocin.
OS   Cyprinus carpio (Common carp).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Cyprinus.
OX   NCBI_TaxID=7962;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Pituitary;
RA   Acher R., Chauvet J., Chauvet M.-T., Crepy D.;
RT   "Characterization of neurohypophyseal hormones from a fresh water bony
RT   fish, the carp (Cyprinus carpio). Comparison with hormones from sea
RT   water bony fishs.";
RL   Comp. Biochem. Physiol. 14:245-254(1965).
CC   -!- FUNCTION: ANTIDIURETIC HORMONE.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   PIR; A61364; A61364.
DR   InterPro; IPR000981; Neurhyp_horm.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Amidation.
FT   DISULFID      1      6
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  969 MW;  17FF476EB455B04B CRC64;

  Query Match             27.1%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLS 3
              |:|
Db          2 YIS 4


RESULT 6
OXYT_RAJCL
ID   OXYT_RAJCL     STANDARD;      PRT;     9 AA.
AC   P42994;
DT   01-NOV-1995 (Rel. 32, Created)
```

```
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Glumitocin.
OS   Raja clavata (Thornback ray).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;
OC   Rajiformes; Rajidae; Raja.
OX   NCBI_TaxID=7781;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=66123415; PubMed=5880565;
RA   Acher R., Chauvet J., Chauvet M.-T., Crepy D.;
RT   "Phylogeny of neurophyophyseal peptides: isolation of a new hormone,
RT   glumitocin (Ser 4-Gln 8-ocytocin) present in a cartilaginous fish,
RT   the ray (Raia clavata).";
RL   Biochim. Biophys. Acta 107:393-396(1965).
CC   -!- FUNCTION: ANTIDIURETIC HORMONE.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   InterPro; IPR000981; Neurhyp_horm.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Amidation.
FT   DISULFID      1      6
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  984 MW;  17E9C76EB455B04B CRC64;

  Query Match             27.1%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLS 3
              |:|
Db          2 YIS 4


RESULT 7
PGLR_DIAAB
ID   PGLR_DIAAB     STANDARD;      PRT;     9 AA.
AC   P81179;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Endo-polygalacturonase (PG) (EC 3.2.1.15) (Fragment).
OS   Diaprepes abbreviatus (Sugarcane rootstalk borer weevil).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC   Phytophaga; Curculionidae; Entiminae; Entimini; Diaprepes.
OX   NCBI_TaxID=13040;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Larval gut;
RA   Doostdar H., McCollum T.G., Mayer R.T.;
RT   "Purification and characterization of an endo-polygalacturonase from
RT   the gut of West Indies sugarcane rootstalk borer weevil (Diaprepes
RT   abbreviatus L.) larvae.";
RL   Comp. Biochem. Physiol. 118B:861-867(1997).
CC   -!- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
CC       galactosiduronic linkages in pectate and other galacturonans.
CC   -!- INDUCTION: INHIBITED BY CITRUS PGIP.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
CC       9.4, ITS MW IS: 44.5 kDa.
CC   -!- SIMILARITY: WEAK, TO OTHER POLYGALACTURONASES.
KW   Hydrolase; Glycosidase; Cell wall.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1041 MW;  1F49087042DB41BB CRC64;

  Query Match             25.0%;  Score 12;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSG 4
              |: |
Db          4 YVIG 7
```

```
RESULT 8
FAR3_HIRME
ID   FAR3_HIRME     STANDARD;      PRT;     4 AA.
AC   P42562;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFamide-like neuropeptide YLRF-amide.
OS   Hirudo medicinalis (Medicinal leech).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92195954; PubMed=1686933;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Peptides 12:897-908(1991).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       4      4       AMIDATION.
SQ   SEQUENCE   4 AA;  598 MW;  69D4073B30000000 CRC64;

  Query Match             22.9%;  Score 11;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YL 2
              ||
Db          1 YL 2


RESULT 9
PRCT_PERAM
ID   PRCT_PERAM     STANDARD;      PRT;     5 AA.
AC   P01373;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   Proctolin.
OS   Periplaneta americana (American cockroach),
OS   Limulus polyphemus (Atlantic horseshoe crab), and
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattoidea;
OC   Blattidae; Periplaneta.
OX   NCBI_TaxID=6978, 6850, 6759;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=P.americana;
RX   MEDLINE=76074708; PubMed=576;
RA   Starratt A.N., Brown B.E.;
RT   "Structure of the pentapeptide proctolin, a proposed neurotransmitter
RT   in insects.";
RL   Life Sci. 17:1253-1256(1975).
RN   [2]
RP   BIOLOGICAL SOURCE.
RC   SPECIES=P.americana;
RX   MEDLINE=81225865; PubMed=6113690;
RA   O'Shea M., Adams M.E.;
RT   "Pentapeptide (proctolin) associated with an identified neuron.";
RL   Science 213:567-569(1981).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=L.polyphemus;
RX   MEDLINE=90287800; PubMed=2356151;
RA   Groome J.R., Tillinghast E.K., Townley M.A., Vetrovs A.,
RA   Watson W.H. III, Hunt D.F., Griffin P.R., Alexander J.E.,
RA   Shabanowitz J.;
RT   "Identification of proctolin in the central nervous system of the
```

```
RT   horseshoe crab, Limulus polyphemus.";
RL   Peptides 11:205-211(1990).
RN   [4]
RP   SEQUENCE.
RC   SPECIES=C.maenas;
RX   MEDLINE=86232789; PubMed=2872661;
RA   Stangier J., Dircksen H., Keller R.;
RT   "Identification and immunocytochemical localization of proctolin in
RT   pericardial organs of the shore crab, Carcinus maenas.";
RL   Peptides 7:67-72(1986).
CC   -!- FUNCTION: STIMULATES CARDIAC OUTPUT AND HINDGUT MOTILITY,
CC       MODULATES VISCERAL AND SKELETAL MUSCLE IN MANY ARTHROPODS.
CC   -!- TISSUE SPECIFICITY: FOUND IN THE LATERAL WHITE NEURONS AND IN
CC       THE CRAB PERICARDIAL ORGANS.
DR   PIR; A01644; HOROHA.
DR   PIR; A60411; A60411.
KW   Neuropeptide.
SQ   SEQUENCE   5 AA;  649 MW;  71B7673B44600000 CRC64;

  Query Match             22.9%;  Score 11;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YL 2
              ||
Db          2 YL 3


RESULT 10
FAR2_ASCSU
ID   FAR2_ASCSU     STANDARD;      PRT;     7 AA.
AC   P31890;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   FMRFamide-like neuropeptide AF2.
OS   Ascaris suum (Pig roundworm) (Ascaris lumbricoides), and
OS   Panagrellus redivivus.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC   Ascarididae; Ascaris.
OX   NCBI_TaxID=6253, 6233;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=A.suum;
RX   MEDLINE=93324431; PubMed=8332542;
RA   Cowden C., Stretton A.O.W.;
RT   "AF2, an Ascaris neuropeptide: isolation, sequence, and bioactivity.";
RL   Peptides 14:423-430(1993).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=P.redivivus;
RX   MEDLINE=95060998; PubMed=7970891;
RA   Maule A.G., Shaw C., Bowman J.W.;
RT   "The FMRFamide-like neuropeptide AF2 (Ascaris suum) is present in the
RT   free-living nematode, Panagrellus redivivus (Nematoda, Rhabditida).";
RL   Parasitology 109:351-356(1994).
CC   -!- FUNCTION: HAS EFFECTS ON MUSCLE TENSION.
CC   -!- TISSUE SPECIFICITY: FOUND IN THE NERVE CORDS AND A VARIETY OF
CC       GANGLIA PARTICULARLY IN THE ANTERIOR REGIONS.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       7      7       AMIDATION.
SQ   SEQUENCE   7 AA;  992 MW;  69D4073B5B11E350 CRC64;

  Query Match             22.9%;  Score 11;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YL 2
              ||
Db          4 YL 5


RESULT 11
GFRP_MOUSE
ID   GFRP_MOUSE     STANDARD;      PRT;     7 AA.
AC   P99025;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   GTP cyclohydrolase I feedback regulatory protein (P35) (Fragment).
GN   GCHFR OR GFRP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Liver;
RA   Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,
RA   Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
RA   Cowthorne M.;
RL   Submitted (AUG-1998) to the SWISS-PROT data bank.
CC   -!- FUNCTION: MEDIATES TETRAHYDROBIOPTERIN INHIBITION OF GTP
CC       CYCLOHYDROLASE I. THIS INHIBITION IS REVERSED BY L-PHENYLALANINE
CC       (BY SIMILARITY).
CC   -!- SUBUNIT: Homodimer (By similarity).
DR   SWISS-2DPAGE; P99025; MOUSE.
FT   INIT_MET      0      0
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  806 MW;  71B5B057273B4700 CRC64;

  Query Match             22.9%;  Score 11;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YL 2
              ||
Db          2 YL 3


RESULT 12
CAD1_ENTFA
ID   CAD1_ENTFA     STANDARD;      PRT;     8 AA.
AC   P13268;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Sex pheromone CAD1.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=85051889; PubMed=6437872;
RA   Mori M., Sagakami Y., Narita M., Isogai A., Fujino M., Kitada C.,
RA   Craig R.A., Clewell D.B., Suzuki A.;
RT   "Isolation and structure of the bacterial sex pheromone, cAD1, that
RT   induces plasmid transfer in Streptococcus faecalis.";
RL   FEBS Lett. 178:97-100(1984).
CC   -!- FUNCTION: CAD1 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC       HEMOLYSIN PLASMID PAD1.
KW   Pheromone.
SQ   SEQUENCE   8 AA;  819 MW;  047DD732C735B9C7 CRC64;

  Query Match             22.9%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |:|
Db          6 LAG 8


RESULT 13
```

```
COW2_CONPU
ID   COW2_CONPU     STANDARD;      PRT;     8 AA.
AC   P58785;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Leu-contryphan-P.
OS   Conus purpurascens (Purple cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC   Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
OX   NCBI_TaxID=41690;
RN   [1]
RP   SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC   STRAIN=Clipperton Island; TISSUE=Venom;
RX   MEDLINE=99388839; PubMed=10461743;
RA   Jacobsen R.B., Jimenez E.C., De la Cruz R.G.C., Gray W.R., Cruz L.J.,
RA   Olivera B.M.;
RT   "A novel D-leucine-containing Conus peptide: diverse conformational
RT   dynamics in the contryphan family.";
RL   J. Pept. Res. 54:93-99(1999).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC   -!- MASS SPECTROMETRY: MW=888.4; METHOD=LSIMS.
CC   -!- SIMILARITY: BELONGS TO THE CONTRYPHAN FAMILY.
KW   Toxin; Hydroxylation; D-amino acid.
FT   DISULFID      2      8
FT   MOD_RES       4      4       D-LEUCINE.
SQ   SEQUENCE   8 AA;  890 MW;  75A367672732CEB8 CRC64;

  Query Match             22.9%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 CLNL 9
              |: |
Db          2 CVLL 5


RESULT 14
CONO_CONST
ID   CONO_CONST     STANDARD;      PRT;     9 AA.
AC   P05487;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Arg-conopressin S.
OS   Conus striatus (Striated cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC   Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
OX   NCBI_TaxID=6493;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=88058932; PubMed=3680228;
RA   Cruz L.J., de Santos V., Zafaralla G.C., Ramilo C.A., Zeikus R.D.,
RA   Gray W.R., Olivera B.M.;
RT   "Invertebrate vasopressin/oxytocin homologs. Characterization of
RT   peptides from Conus geographus and Conus straitus venoms.";
RL   J. Biol. Chem. 262:15821-15824(1987).
RN   [2]
RP   REVIEW.
RX   MEDLINE=89024586; PubMed=3052286;
RA   Gray W.R., Olivera B.M., Cruz L.J.;
RT   "Peptide toxins from venomous Conus snails.";
RL   Annu. Rev. Biochem. 57:665-700(1988).
CC   -!- FUNCTION: Targets vasopressin-oxytocin related receptors.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   PIR; B28495; B28495.
DR   InterPro; IPR000981; Neurhyp_horm.
DR   Pfam; PF00220; hormone4; 1.
```

```
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Amidation.
FT   DISULFID      1      6
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1031 MW;  17EB176EB4540050 CRC64;

  Query Match             22.9%;  Score 11;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 CL 7
              |:
Db          1 CI 2


RESULT 15
MOSF_CLYJA
ID   MOSF_CLYJA     STANDARD;      PRT;     9 AA.
AC   P19853;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   [Phe-6]-mosact.
OS   Clypeaster japonicus (Sand dollar).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Gnathostomata; Clypeasteroida;
OC   Clypeasteridae; Clypeaster.
OX   NCBI_TaxID=7644;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Egg jelly;
RA   Suzuki N., Kurita M., Yoshino K.I., Kajiura H., Nomura K.,
RA   Yamaguchi M.;
RT   "Purification and structure of mosact and its derivatives from the
RT   egg jelly of the sea urchin Clypeaster japonicus.";
RL   Zool. Sci. 4:649-656(1987).
CC   -!- FUNCTION: Stimulates sperm respiration and motility.
DR   PIR; JN0027; JN0027.
SQ   SEQUENCE   9 AA;  924 MW;  93245729CDC5BAB5 CRC64;

  Query Match             22.9%;  Score 11;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSG 4
              :| |
Db          6 FLIG 9


Search completed: January 12, 2004, 14:29:03
Job time : 6.25 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         January 12, 2004, 14:21:08 ; Search time 32.25 Seconds
                                           (without alignments)
                                           44.296 Million cell updates/sec

Title:          US-09-529-121A-2
Perfect score:  45
Sequence:       1 YLSGADLNL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters:      179625

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_19Jun03:*
                1:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
                2:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
                3:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
                4:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
                5:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
                6:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
                7:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
                8:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
                9:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
                10:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
                11:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
                12:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
                13:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                14:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
                15:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
                16:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                17:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
                18:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                19:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
                20:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                21:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
                22:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
                23:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
                24:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES

                      %
Result              Query
  No.    Score    Match Length DB  ID                           Description
-------------------------------------------------------------------------------
    1      45    100.0      9  20  AAY09526                     Carcinoembryonic a
    2      45    100.0      9  21  AAB13750                     Peptide fragment #
    3      45    100.0      9  22  AAE05124                     Modified carcinoem
    4      45    100.0      9  22  AAB97818                     Carcinoembryonic a
    5      45    100.0      9  23  AAB47917                     Modified CEA epito
    6      45    100.0      9  23  AAE19089                     HLA-A24 restricted
    7      43     95.6      9  20  AAY09527                     Carcinoembryonic a
    8      40     88.9      9  18  AAW39723                     Human carcina-embr
    9      40     88.9      9  19  AAW77134                     CEA synthetic pept
```

```
   10      40     88.9      9  19  AAW70045                     CEA derived HLA-A2
   11      40     88.9      9  20  AAY47655                     Immunogenic peptid
   12      40     88.9      9  20  AAY09525                     Carcinoembryonic a
   13      40     88.9      9  21  AAB13749                     Peptide fragment #
   14      40     88.9      9  22  AAB82776                     Carcinoembryonic a
   15      40     88.9      9  22  AAE05123                     Carcinoembryonic a
   16      40     88.9      9  22  AAE02673                     Human CEA epitopic
   17      40     88.9      9  22  AAE00463                     Human tumour CEA e
   18      40     88.9      9  23  AAE26805                     Human HLA-A2.1 res
   19      40     88.9      9  23  ABG79073                     Human CEA class I
   20      40     88.9      9  23  AAU95893                     Immunogenic peptid
   21      40     88.9      9  23  AAE19088                     HLA-A24 restricted
   22      38     84.4      9  20  AAY09528                     Carcinoembryonic a
   23      37     82.2      9  21  AAY54173                     HLA binding peptid
   24      37     82.2      9  22  AAU26560                     Human Leukocyte An
   25      37     82.2      9  22  AAB99681                     HLA A2 binding CTL
   26      36     80.0      9  20  AAY09529                     Carcinoembryonic a
   27      34     75.6      9  22  AAB75854                     Tumour associated
   28      33     73.3      9  17  AAW00680                     Peptide comprising
   29      28     62.2      9  23  AAU82064                     CEA antigenic pept
   30      25     55.6      7  20  AAY41847                     Rheumatoid arthrit
   31      25     55.6      9  18  AAW38383                     Synthetic pMEL17 p
   32      25     55.6      9  20  AAY47062                     Immunogenic peptid
   33      25     55.6      9  23  AAE31162                     Human sp100 peptid
   34      25     55.6      9  23  AAE31390                     Human PM17 peptide
   35      24     53.3      6  23  AAE24489                     BoNT/E N-terminal
   36      24     53.3      9  17  AAW00690                     NCA analogue of re
   37      24     53.3      9  19  AAW70078                     B. stearothermophi
   38      23     51.1      7  20  AAY41846                     Rheumatoid arthrit
   39      23     51.1      9  11  AAR07966                     Tryptic fragment T
   40      23     51.1      9  22  AAU23920                     Human MHC class I
   41      23     51.1      9  22  AAU24043                     Human MHC class I
   42      23     51.1      9  22  AAU24335                     Human MHC class I
   43      23     51.1      9  22  AAB88248                     Hsp-65 peptide epi
   44      23     51.1      9  24  ABJ38049                     Human cytomegalovi
   45      23     51.1      9  24  ABR16172                     Human cancer-relat

                              ALIGNMENTS

RESULT 1
AAY09526
ID   AAY09526 standard; peptide; 9 AA.
XX
AC   AAY09526;
XX
DT   20-JUL-1999  (first entry)
XX
DE   Carcinoembryonic antigen peptide agonist SEQ ID NO:2.
XX
KW   Carcinoembryonic antigen; CEA; human; agonist; antagonist;
KW   immune response; carcinoma; gastrointestinal; breast; pancreatic;
KW   bladder; ovarian; lung; prostatic; T cell proliferation; cancer;
KW   adoptive transfer therapy; autoimmune reaction; immunotherapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9919478-A1.
XX
PD   22-APR-1999.
XX
PF   22-SEP-1998;   98WO-US19794.
XX
PR   10-OCT-1997;   97US-0061589.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Barzaga E,  Schlom J,  Zaremba S;
XX
DR   WPI; 1999-326544/27.
XX
```

```
PT   Peptide agonists and antagonists of carcinoembryonal antigen
XX
PS   Claim 5; Page 53; 72pp; English.
XX
CC   The present invention describes peptides (A) that comprise agonists (Ia)
CC   or antagonists (Ib) of human carcinoembryonal antigen (CEA). (Ia) are
CC   used in vaccines to kill or inhibit carcinoma cells that express CEA or
CC   its epitopes, particularly for treating gastrointestinal, breast,
CC   pancreatic, bladder, ovarian, lung or prostatic carcinoma. They can also
CC   be used to proliferate T cells, e.g. from vaccinated subjects, for use
CC   in adoptive transfer therapy. (Ib) are used to inhibit CEA-specific
CC   immune responses, e.g. in vaccinated subjects, to prevent an autoimmune
CC   reaction to cancer immunotherapy (i.e. to prevent attack on normal but
CC   CEA-expressing cells). (Ia) are more active than native sequence (I) and
CC   generate a highly specific and systemic anti-CEA response. Cytotoxic T
CC   cells generated recognize both (Ia) and native CEA epitopes. The present
CC   sequence represents a specifically claimed example of (Ia).
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 45;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADLNL 9
              |||||||||
Db          1 YLSGADLNL 9


RESULT 2
AAB13750
ID   AAB13750 standard; peptide; 9 AA.
XX
AC   AAB13750;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Peptide fragment # 2 from human CEA.
XX
KW   Human; T-cell; immune response; antigen; epitope; B7 family molecule;
KW   Leukocyte function-associated antigen-3; LFA-3;
KW   Intercellular adhesion molecule-1; ICAM-1; vaccine; immunotherapy;
KW   colon polyp; Crohn's disease; ulcerative colitis; breast lesion;
KW   tumour; CEA.
XX
OS   Homo sapiens.
XX
PN   WO200034494-A1.
XX
PD   15-JUN-2000.
XX
PF   12-NOV-1999;   99WO-US26866.
XX
PR   09-DEC-1998;   98US-0111582.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (THER-) THERION BIOLOGICS CORP.
XX
PI   Schlom J,  Hodge J,  Panicali D;
XX
DR   WPI; 2000-431307/37.
XX
PT   Novel recombinant vector useful as immunogens and vaccines for
PT   stimulating and enhancing immunological responses to target cells and
PT   antigens expresses multiple co-stimulatory molecules such as B7-1,
PT   LFA-3, ICAM-1  -
XX
PS   Claim 18; Page 35; 188pp; English.
XX
CC   Costimulatory molecules have important roles in T-cell activation and
CC   therefore the immune response. The present invention relates to
CC   recombinant vectors which comprise of foreign nucleic acid sequences
CC   encoding at least three costimulatory molecules: a B7 family molecule,
CC   Leukocyte function-associated antigen-3 (LFA-3, human CD58) and
CC   Intercellular adhesion molecule-1 (ICAM-1, CD54) and optionally a foreign
CC   gene encoding a target antigen or immunological epitope. The present
CC   sequence is one such target antigen used in the present invention. The
CC   present sequence is a tumour-associated antigen. The vector of the
CC   present invention would be useful for providing an enhanced immune
CC   response to the present target antigen. The vector of the present
CC   invention may therefore be useful in immunotherapy for treating or
CC   preventing diseases caused by viruses, bacteria, protozoans, parasites,
CC   premalignant cells and tumour cells. The recombinant vector can be used
CC   to treat or prevent preneoplastic or hyperplastic states such as colon
CC   polyps, Crohn's disease, ulcerative colitis and breast lesions.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 45;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADLNL 9
              |||||||||
Db          1 YLSGADLNL 9


RESULT 3
AAE05124
ID   AAE05124 standard; peptide; 9 AA.
XX
AC   AAE05124;
XX
DT   18-SEP-2001  (first entry)
XX
DE   Modified carcinoembryonic antigen (CEA) peptide, CAP-6D.
XX
KW   Tumour-associated antigen; TAA; cytostatic; vaccine; gene therapy;
KW   immune response; tetanus toxoid; TT; diphtheria toxoid; DT; prophylactic;
KW   cancer; therapeutic; carcinoembryonic antigen; CEA.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 6
FT                   /note= "Wild type Asn substituted with Asp"
XX
PN   WO200149317-A2.
XX
PD   12-JUL-2001.
XX
PF   05-JAN-2001; 2001WO-CA00005.
XX
PR   05-JAN-2000; 2000US-0174587.
XX
PA   (AVET ) AVENTIS PASTEUR LTD.
XX
PI   Emtage P,  Barber BH,  Sambhara S,  Sia CDY;
XX
DR   WPI; 2001-441790/47.
XX
PT   Enhancing immune response to antigen such as tumor antigen for treating
PT   cancer in an animal involves administering an inducing agent to the
PT   animal followed by administering inducing agent-antigen mixture  -
XX
PS   Example 2; Page 31; 62pp; English.
XX
CC   The invention relates to a method of enhancing an immune response against
CC   tumour-associated antigens (TAAs), such as GP100 and carcinoembryonic
CC   antigen (CEA) in an animal. The method involves priming of the animal
CC   with an inducing agent such as tetanus toxoid (TT) or diphtheria toxoid
CC   (DT), subsequently followed by administration of an inducing agent-
CC   antigen mixture. The method provides the enhancement or augmentation of
CC   the immune response to the antigen and/or improves a vaccination protocol
CC   by allowing use of less antigen. The immunisation of the animal with
CC   tumour-associated antigen is useful for the prophylactic or therapeutic
```

```
CC   treatment of cancer. The present sequence is modified carcinoembryonic
CC   antigen (CEA) peptide fragment related to the invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 45;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADLNL 9
              |||||||||
Db          1 YLSGADLNL 9


RESULT 4
AAB97818
ID   AAB97818 standard; Peptide; 9 AA.
XX
AC   AAB97818;
XX
DT   08-AUG-2001  (first entry)
XX
DE   Carcinoembryonic antigen (CEA) modified antigen SEQ ID NO:113.
XX
KW   Virus; adenovirus; poxvirus; alphavirus; immune response; gp100;
KW   tumour antigen; CEA; carcinoembryonic antigen; immunostimulant;
KW   cytostatic; immunotherapy; interferon-gamma; IFN-gamma; cancer.
XX
OS   Unidentified.
XX
PN   WO200130382-A1.
XX
PD   03-MAY-2001.
XX
PF   20-OCT-2000; 2000WO-CA01253.
XX
PR   22-OCT-1999;   99US-0160879.
PR   07-AUG-2000; 2000US-0223325.
XX
PA   (AVET ) AVENTIS PASTEUR LTD.
XX
PI   Berinstein N,  Tartaglia J,  Moingeon P,  Barber B;
XX
DR   WPI; 2001-308587/32.
XX
PT   Inducing immune response to tumor antigen, useful in immunotherapy of
PT   cancer, by administering the antigen to a lymphatic site  -
XX
PS   Claim 19; Page 9; 60pp; English.
XX
CC   The present invention describes a method for inducing an immune response,
CC   in an animal, to a tumour antigen (Ag) comprising administering Ag, or
CC   nucleic acid (I) that encodes it, to a lymphatic site. Cynomolgus monkeys
CC   (Macaca fascicularis) were injected with a modified form of gp100 antigen
CC   (a) into the left inguinal lymph node or (b) subcutaneously. Both animals
CC   of (a) developed a cell-mediated response (indicated by production of
CC   interferon-gamma from T lymphocytes when exposed to gp100 peptides), but
CC   only 2 of 4 animals of (b) did so. Also animals in (a) produced a far
CC   greater antibody response to gp100. The method is used in immunotherapy
CC   of a wide range of cancers through induction of a specific immune
CC   response (humoral and cellular) against the tumour antigens. When
CC   administered to a lymphatic site, Ag (or (I)) induces a stronger immune
CC   response than administration by other routes and may also break tolerance
CC   to Ag. AAB97708 and AAB97709 represent gp100 epitopes; AAB97710 to
CC   AAB97815 represent peptides derived from gp100 which stimulate interferon
CC   (IFN)-gamma production; AAH20120 encodes the modified gp100 protein given
CC   in AAB97816; AAH20121 encodes the modified carcinoembryonic antigen (CEA)
CC   protein given in AAB97817; and AAB97818 represents a CEA modified antigen
CC   peptide, all of which are used in the exemplification of the present
CC   invention.
XX
SQ   Sequence   9 AA;
```

```
  Query Match             100.0%;  Score 45;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADLNL 9
              |||||||||
Db          1 YLSGADLNL 9

RESULT 5
AAB47917
ID   AAB47917 standard; peptide; 9 AA.
XX
AC   AAB47917;
XX
DT   16-MAY-2002  (first entry)
XX
DE   Modified CEA epitope, CEA(6D).
XX
KW   CAP-1; epitope; carcinoembryonic antigen; CEA; agonist; immune response;
KW   carcinoma; gastrointestinal; breast; pancreatic; bladder; ovarian;
KW   lung; prostate; cancer.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 6
FT                   /label= N6D
XX
PN   WO200210379-A2.
XX
PD   07-FEB-2002.
XX
PF   27-JUL-2001; 2001WO-CA01092.
XX
PR   31-JUL-2000; 2000US-222043P.
XX
PA   (AVET ) AVENTIS PASTEUR LTD.
PA   (THER-) THERION BIOLOGICS.
PA   (USSH ) US NAT CANCER INST.
XX
PI   Berinstein N,  Tartaglia J,  Tine JA,  Panicali DL,  Gritz L;
PI   Schlom J;
XX
DR   WPI; 2002-206189/26.
XX
PT   Carcinoembryonic antigen agonist polypeptide for inducing an immune
PT   response in animal against antigen and for inhibiting an epitope
PT   antigen expressing carcinoma cell, comprises a modified antigen epitope
PT   -
XX
PS   Claim 1; Page 38; 69pp; English.
XX
CC   This sequence represents a modified CAP-1 epitope of carcinoembryonic
CC   antigen (CEA) which was used as part of the CEA agonist polypeptide of
CC   the invention.  The modification of position 6 of this peptide from Asp
CC   to Asn increases its immunogenicity.  The CEA agonist polypeptide of
CC   the invention, or DNA encoding it, are useful for:
CC   (i) inducing an immune response in an animal directed against a CEA
CC   protein or fragment, CEA agonist, a CEA epitope, a modified CEA epitope,
CC   cells expressing or binding a CEA protein or fragment; and
CC   (ii) inhibiting a CEA epitope expressing carcinoma cell, which is a
CC   gastrointestinal, breast, pancreatic, bladder, ovarian, lung or
CC   prostate carcinoma cell in a patient, hence is useful for manufacture
CC   of a medicament for the treatment of cancer.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 45;  DB 23;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADLNL 9
```

```
                |||||||||
Db            1 YLSGADLNL 9

RESULT 6
AAE19089
ID   AAE19089 standard; peptide; 9 AA.
XX
AC   AAE19089;
XX
DT   21-MAY-2002  (first entry)
XX
DE   HLA-A24 restricted target antigen CEA immunological epitope #3.
XX
KW   Human leukocyte antigen; HLA; pharmaceutical composition; target antigen;
KW   immunological epitope; replication-defective virus; RDV; immune response;
KW   chemotherapy; granulocyte-monocyte-colony stimulating factor; cytostatic;
KW   GM-CSF; MHC; major histocompatibility complex; tumour; head; pancreatic;
KW   neck; breast; prostate; colorectal; melanoma; myeloidysplastic syndrome;
KW   metastatic breast skin lesion; corticosteroid therapy; erythropoietin;
KW   cytopenia; neutropenia; vaccine; immunostimulant.
XX
OS   Homo sapiens.
XX
PN   WO200195919-A2.
XX
PD   20-DEC-2001.
XX
PF   15-JUN-2001; 2001WO-US19201.
XX
PR   15-JUN-2000; 2000US-211717P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (THER-) THERION BIOLOGICS CORP.
XX
PI   Schlom J,  Greiner JW,  Kass E,  Panicali D;
XX
DR   WPI; 2002-205852/26.
XX
PT   Composition for enhancing immune responses, particularly anti-tumor
PT   responses and treating neutropenia, cytopenia, comprises
PT   replication-defective virus encoding granulocyte-monocyte-colony
PT   stimulating factor  -
XX
PS   Claim 9; Page 15; 118pp; English.
XX
CC   The present invention relates to a pharmaceutical composition comprising
CC   a replication-defective virus (RDV) encoding granulocyte-monocyte-colony
CC   stimulating factor (GM-CSF). The invention is useful for enhancing cell-
CC   mediated or humoral immune response in an individual, by enhancing
CC   migration of APC expressing CD11c^+/I-Ab^+, major histocompatibility
CC   complex (MHC) class II, at an injection site, regional lymph node at a
CC   tumour site, APC proliferation or function, CD4^+T or CD8^+T cell
CC   activation, interleukin (IL)-2, interferon (IFN)-gamma or tumour necrosis
CC   factor (TNF)-alpha production or their combinations. The composition
CC   enhances an antigen-specific T-cell response in an individual to a target
CC   antigen or its immunological epitope and an anti-tumour response in an
CC   individual with a head tumour, neck, breast, pancreatic, prostate,
CC   colorectal or metastatic tumour or melanoma, or metastatic breast skin
CC   lesion. The invention is further useful for treating neutropenia
CC   resulting from chemotherapy, corticosteroid therapy, irradiation or an
CC   infection, by raising the neutrophil count to normal levels and for
CC   treating cytopenias in patients with myeloidysplastic syndrome in
CC   combination with erythropoietin, by increasing neutrophil count and
CC   erythroid precursors. The composition enhances immune response to
CC   vaccines such as DPT, Td, DtaP, Hib, DtaP-Hib, MMR, Hepatitis A,
CC   hepatitis B, lyme's disease, influenza, tetravalent meningococcal
CC   polysaccharide, pneumococcal polysaccharide, anthrax, cholera, plague,
CC   yellow fever and Bacillus Calmette-Guerin vaccine. The present sequence
CC   is human leukocyte antigen (HLA)-restricted target tumour antigen
CC   immunological epitope.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 45;  DB 23;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy            1 YLSGADLNL 9
                |||||||||
Db            1 YLSGADLNL 9

RESULT 7
AAY09527
ID   AAY09527 standard; peptide; 9 AA.
XX
AC   AAY09527;
XX
DT   20-JUL-1999  (first entry)
XX
DE   Carcinoembryonic antigen peptide agonist SEQ ID NO:3.
XX
KW   Carcinoembryonic antigen; CEA; human; agonist; antagonist;
KW   immune response; carcinoma; gastrointestinal; breast; pancreatic;
KW   bladder; ovarian; lung; prostatic; T cell proliferation; cancer;
KW   adoptive transfer therapy; autoimmune reaction; immunotherapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9919478-A1.
XX
PD   22-APR-1999.
XX
PF   22-SEP-1998;   98WO-US19794.
XX
PR   10-OCT-1997;   97US-0061589.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Barzaga E,  Schlom J,  Zaremba S;
XX
DR   WPI; 1999-326544/27.
XX
PT   Peptide agonists and antagonists of carcinoembryonal antigen
XX
PS   Claim 5; Page 53; 72pp; English.
XX
CC   The present invention describes peptides (A) that comprise agonists (Ia)
CC   or antagonists (Ib) of human carcinoembryonal antigen (CEA). (Ia) are
CC   used in vaccines to kill or inhibit carcinoma cells that express CEA or
CC   its epitopes, particularly for treating gastrointestinal, breast,
CC   pancreatic, bladder, ovarian, lung or prostatic carcinoma. They can also
CC   be used to proliferate T cells, e.g. from vaccinated subjects, for use
CC   in adoptive transfer therapy. (Ib) are used to inhibit CEA-specific
CC   immune responses, e.g. in vaccinated subjects, to prevent an autoimmune
CC   reaction to cancer immunotherapy (i.e. to prevent attack on normal but
CC   CEA-expressing cells). (Ia) are more active than native sequence (I) and
CC   generate a highly specific and systemic anti-CEA response. Cytotoxic T
CC   cells generated recognize both (Ia) and native CEA epitopes. The present
CC   sequence represents a specifically claimed example of (Ia).
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 20;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy            1 YLSGADLNL 9
                ||||||:||
Db            1 YLSGADINL 9

RESULT 8
AAW39723
```

```
ID   AAW39723 standard; peptide; 9 AA.
XX
AC   AAW39723;
XX
DT   11-JUN-1998  (first entry)
XX
DE   Human carcina-embryonic antigen (CEA) peptide (pos. 571-579).
XX
KW   T cell epitope; immune response; human leukocyte antigen; HLA Class I;
KW   vaccine; immunogenic; major histocompatablilty complex; MHC; B cell;
KW   disease; anti-tumour; anti-viral.
XX
OS   Homo sapiens.
XX
PN   WO9741440-A1.
XX
PD   06-NOV-1997.
XX
PF   28-APR-1997;   97WO-NL00229.
XX
PR   23-DEC-1996;   96EP-0203670.
PR   26-APR-1996;   96EP-0201145.
XX
PA   (UYLE-) RIJKSUNIV LEIDEN.
PA   (SCIS-) SCI SEED CAPITAL INVESTMENTS BV.
XX
PI   Kast WM,  Melief CJM,  Offringa R,  Toes REM,  Van Der Burg SH;
XX
DR   WPI; 1997-549891/50.
XX
PT   Method of selecting T cell peptide epitope(s) - by measuring the
PT   stability of HLA class I-peptide complexes on intact B cells
XX
PS   Example 3; Page 85; 109pp; English.
XX
CC   Peptides AAW39430-W39734 are used in a novel method for the selection of
CC   immunogenic T-cell peptide epitopes present in polypeptide antigens. The
CC   method involves the identification of peptide sequences capable of
CC   binding to an HLA (human leukocyte antigen) class I molecule and
CC   measuring the binding of this epitope peptide to the HLA class I
CC   peptide. The stability of binding of the peptide and MHC (major
CC   histocompatablilty complex) class I molecule is measured on intact human
CC   B cells carrying the MHC molecule at their cell surfaces. The method can
CC   be used to select peptide epitopes for generating vaccines against a
CC   disease associated with the polypeptide, e.g. cancers or AIDS. The
CC   peptide epitopes are especially T-cell peptide epitopes with strong
CC   anti-tumour and anti-viral immune responses. Peptide AAW39723 is derived
CC   from the human carcino-embryonic antigen (CEA) and has the ability to
CC   bind to the human MHC Class I allele HLA-A2.1.
XX
SQ   Sequence   9 AA;

  Query Match             88.9%;  Score 40;  DB 18;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADLNL 9
                |||||:|||
Db          1 YLSGANLNL 9

RESULT 9
AAW77134
ID   AAW77134 standard; peptide; 9 AA.
XX
AC   AAW77134;
XX
DT   16-NOV-1998  (first entry)
XX
DE   CEA synthetic peptide epitope 1.
XX
KW   Tyrosinase; tyrosinase cytotoxic lymphocyte response;
KW   cytotoxic T lymphocyte; cysteine-depleted; melanoma.
```

```
XX
OS   Synthetic.
XX
PN   WO9833810-A2.
XX
PD   06-AUG-1998.
XX
PF   29-JAN-1998;   98WO-US01592.
XX
PR   30-JAN-1997;   97US-0037781.
XX
PA   (UYVI-) UNIV VIRGINIA PATENT FOUND.
XX
PI   Engelhard VH,  Hunt DF,  Kittlesen D,  Slingluff CL;
XX
DR   WPI; 1998-437388/37.
XX
PT   Disease specific immunogen  - comprises disease specific cytotoxic T
PT   lymphocyte epitope used to elicit melanoma specific CTL response
XX
PS   Disclosure; Page 27; 93pp; English.
XX
CC   The peptide epitope AAW77119-W77138 were created for human
CC   tumour-specific cytotoxic T lymphocyte response.  These peptides are are
CC   cysteine- depleted mutants of a native disease-specific CTL epitope.  The
CC   cysteine- depleted CTL epitopes elicit a stronger or more specific CTL
CC   response than the native epitope.  The epitopes can be used in a
CC   disease-specific immunogen to protect a mammal against disease in
CC   particular melanomas. The peptides may also be used to screen a sample
CC   for the presence of an antigen with the same epitope, or with a different
CC   cross-reactive epitope.
XX
SQ   Sequence   9 AA;

  Query Match             88.9%;  Score 40;  DB 19;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADLNL 9
                |||||:|||
Db          1 YLSGANLNL 9

RESULT 10
AAW70045
ID   AAW70045 standard; peptide; 9 AA.
XX
AC   AAW70045;
XX
DT   22-OCT-1998  (first entry)
XX
DE   CEA derived HLA-A2.1 binding peptide 2 (residues 605-613).
XX
KW   Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW   human leukocyte antigen; HLA; tumour associated antigen; cancer;
KW   antigen presenting cell; APC; immunogenic peptide; immune disorder;
KW   viral infection; AIDS; hepatitis; bacterial infection; malaria; CEA;
KW   fungal infection; tuberculosis; melanoma; carcinoembryonic antigen.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9833888-A1.
XX
PD   06-AUG-1998.
XX
PF   30-JAN-1998;   98WO-US01959.
XX
PR   31-JAN-1997;   97US-0036696.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Celis E,  Sette A,  Sidney J,  Southwood S,  Tsai V;
```

```
XX
DR   WPI; 1998-437445/37.
XX
PT   Production of antigen-specific cytotoxic T cells - by incubating
PT   immunogenic peptide(s) from antigen that binds class I major
PT   histocompatibility complex molecules with pre-treated antigen
PT   presenting cells
XX
PS   Example 6; Page 75; 104pp; English.
XX
CC   Sequences shown in AAW70044 to AAW70052 represent peptides derived from
CC   carcinoembryonic antigen (CEA). The peptides can bind to a human
CC   leukocyte antigen (HLA), HLA-A2.1 and are used to exemplify the method
CC   of invention of producing antigen-specific cytotoxic T cells (CTLs) in
CC   vitro. The method comprises contacting immunogenic peptides from an
CC   antigen that binds class I major histocompatibility complex (MHC)
CC   molecules with antigen presenting cells (APCs) pretreated with
CC   pretreatment growth factors, and incubating the APCs with purified CD8
CC   cells in the presence of at least 2 incubation growth factors, thereby
CC   producing antigen-specific CTLs. A method for specifically killing
CC   target cells in a human patient is also provided which comprises
CC   obtaining a fluid sample containing CTLs from a patient, contacting the
CC   cytotoxic T cells with APCs pretreated with pre-treatment growth
CC   factors, where the APCs comprise class I MHC molecules. The pretreated
CC   APCs are incubated with the cytotoxic growth factors, thereby producing
CC   activated CTLs which are contacted with a carrier to form a composition.
CC   The composition can then be administered to the patient. The activated
CC   CTLs can be used for treating cancers, immune disorders, viral
CC   infections, AIDS, hepatitis, bacterial infection, fungal infection,
CC   malaria or tuberculosis.
XX
SQ   Sequence   9 AA;

  Query Match             88.9%;  Score 40;  DB 19;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADLNL 9
              |||||:|||
Db          1 YLSGANLNL 9

RESULT 11
AAY47655
ID   AAY47655 standard; Peptide; 9 AA.
XX
AC   AAY47655;
XX
DT   01-DEC-1999  (first entry)
XX
DE   Immunogenic peptide having a human leukocyte antigen binding motif #2266.
XX
KW   Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW   immune response; T cell activation; major histocompatibility complex;
KW   cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW   prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW   vaccine; immunisation.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9945954-A1.
XX
PD   16-SEP-1999.
XX
PF   13-MAR-1998;   98WO-US05039.
XX
PR   13-MAR-1998;   98WO-US05039.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Kubo RT,  Sidney J,  Celis E,  Grey HM,  Southwood S;
XX
```

```
DR   WPI; 1999-551214/46.
XX
PT   New immunogenic peptides with HLA binding motif, useful in treatment
PT   and diagnosis of cancers and viral diseases  -
XX
PS   Claim 1; Page 118; 150pp; English.
XX
CC   AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
CC   having a human major histocompatibility complex (MHC) Class I (also
CC   known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC   peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC   HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC   response against the antigen from which the peptide is derived.
CC   Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC   normally induced by an antigen in the form of a peptide fragment bound
CC   to a HLA molecule, rather than the intact foreign antigen itself, and
CC   are particularly important in tumour rejection and in fighting viral
CC   infections. The peptides are therefore useful therapeutically to treat
CC   or prevent viral infections and cancers in mammals (especially humans)
CC   e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC   They can be administered as vaccines to elicit an immune response in
CC   individuals susceptible or otherwise at risk of viral infection or
CC   cancer, or used to treat chronic or acute conditions. They are also
CC   useful diagnostically, and can be used to induce a cytotoxic T cell
CC   response, by contacting a cytotoxic T cell with the peptide e.g. to
CC   produce CTLs ex vivo for infusion back into a patient. The
CC   polynucleotides encoding the immunogenic peptides are also useful
CC   therapeutically and for immunisation as above.
XX
SQ   Sequence   9 AA;

  Query Match             88.9%;  Score 40;  DB 20;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADLNL 9
              |||||:|||
Db          1 YLSGANLNL 9

RESULT 12
AAY09525
ID   AAY09525 standard; peptide; 9 AA.
XX
AC   AAY09525;
XX
DT   20-JUL-1999  (first entry)
XX
DE   Carcinoembryonic antigen peptide agonist CAP-1.
XX
KW   Carcinoembryonic antigen; CEA; human; agonist; antagonist;
KW   immune response; carcinoma; gastrointestinal; breast; pancreatic;
KW   bladder; ovarian; lung; prostatic; T cell proliferation; cancer;
KW   adoptive transfer therapy; autoimmune reaction; immunotherapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9919478-A1.
XX
PD   22-APR-1999.
XX
PF   22-SEP-1998;   98WO-US19794.
XX
PR   10-OCT-1997;   97US-0061589.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Barzaga E,  Schlom J,  Zaremba S;
XX
DR   WPI; 1999-326544/27.
XX
PT   Peptide agonists and antagonists of carcinoembryonal antigen
```

```
XX
PS   Claim 1; Page 53; 72pp; English.
XX
CC   The present invention describes peptides (A) that comprise agonists (Ia)
CC   or antagonists (Ib) of human carcinoembryonal antigen (CEA). (Ia) are
CC   used in vaccines to kill or inhibit carcinoma cells that express CEA or
CC   its epitopes, particularly for treating gastrointestinal, breast,
CC   pancreatic, bladder, ovarian, lung or prostatic carcinoma. They can also
CC   be used to proliferate T cells, e.g. from vaccinated subjects, for use
CC   in adoptive transfer therapy. (Ib) are used to inhibit CEA-specific
CC   immune responses, e.g. in vaccinated subjects, to prevent an autoimmune
CC   reaction to cancer immunotherapy (i.e. to prevent attack on normal but
CC   CEA-expressing cells). (Ia) are more active than native sequence (I) and
CC   generate a highly specific and systemic anti-CEA response. Cytotoxic T
CC   cells generated recognize both (Ia) and native CEA epitopes. The present
CC   sequence represents a specifically claimed example of (Ia).
XX
SQ   Sequence   9 AA;

  Query Match             88.9%;  Score 40;  DB 20;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADLNL 9
              |||||:|||
Db          1 YLSGANLNL 9


RESULT 13
AAB13749
ID   AAB13749 standard; peptide; 9 AA.
XX
AC   AAB13749;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Peptide fragment # 1 from human CEA.
XX
KW   Human; T-cell; immune response; antigen; epitope; B7 family molecule;
KW   Leukocyte function-associated antigen-3; LFA-3;
KW   Intercellular adhesion molecule-1; ICAM-1; vaccine; immunotherapy;
KW   colon polyp; Crohn's disease; ulcerative colitis; breast lesion;
KW   tumour; CEA.
XX
OS   Homo sapiens.
XX
PN   WO200034494-A1.
XX
PD   15-JUN-2000.
XX
PF   12-NOV-1999;   99WO-US26866.
XX
PR   09-DEC-1998;   98US-0111582.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (THER-) THERION BIOLOGICS CORP.
XX
PI   Schlom J,  Hodge J,  Panicali D;
XX
DR   WPI; 2000-431307/37.
XX
PT   Novel recombinant vector useful as immunogens and vaccines for
PT   stimulating and enhancing immunological responses to target cells and
PT   antigens expresses multiple co-stimulatory molecules such as B7-1,
PT   LFA-3, ICAM-1  -
XX
PS   Claim 18; Page 35; 188pp; English.
XX
CC   Costimulatory molecules have important roles in T-cell activation and
CC   therefore the immune response. The present invention relates to
CC   recombinant vectors which comprise of foreign nucleic acid sequences
CC   encoding at least three costimulatory molecules: a B7 family molecule,
CC   Leukocyte function-associated antigen-3 (LFA-3, human CD58) and
```

```
CC   Intercellular adhesion molecule-1 (ICAM-1, CD54) and optionally a foreign
CC   gene encoding a target antigen or immunological epitope. The present
CC   sequence is one such target antigen used in the present invention. The
CC   present sequence is a tumour-associated antigen. The vector of the
CC   present invention would be useful for providing an enhanced immune
CC   response to the present target antigen. The vector of the present
CC   invention may therefore be useful in immunotherapy for treating or
CC   preventing diseases caused by viruses, bacteria, protozoans, parasites,
CC   premalignant cells and tumour cells. The recombinant vector can be used
CC   to treat or prevent preneoplastic or hyperplastic states such as colon
CC   polyps, Crohn's disease, ulcerative colitis and breast lesions.
XX
SQ   Sequence   9 AA;

  Query Match             88.9%;  Score 40;  DB 21;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADLNL 9
              |||||:|||
Db          1 YLSGANLNL 9


RESULT 14
AAB82776
ID   AAB82776 standard; Protein; 9 AA.
XX
AC   AAB82776;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Carcinoembryonic antigen peptide.
XX
KW   Telomerase reverse transcriptase; hTRT; human;
KW   cytotoxic T lymphocyte; major histocompatibility complex; cancer;
KW   tumour; human leucocyte antigen; HLA-A2.1; vaccine;
KW   carcinoembryonic antigen.
XX
OS   Homo sapiens.
XX
PN   WO200160391-A1.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05143.
XX
PR   15-FEB-2000; 2000US-0182685.
PR   15-FEB-2001; 2001US-0182685.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Zanetti M;
XX
DR   WPI; 2001-536552/59.
XX
PT   Vaccine for initiating and enhancing a cytotoxic T lymphocyte response,
PT   for treating cancers or tumours or for inducing immune response against
PT   tumours, comprises a telomerase reverse transcriptase peptide    -
XX
PS   Example 1; Page 12; 52pp; English.
XX
CC   The present sequence is that of a carcinoembryonic antigen peptide
CC   comprising amino acid residues 571-579.  The peptide was used as a
CC   reference peptide in comparison with human telomerase reverse
CC   transcriptase (hTRT) HLA-A2.1+ restricted peptide p540 (see
CC   AAB82772) in a HLA-A2.1 binding/stabilisation assay.  The induction
CC   of CTL responses in vitro and in vivo, and the susceptibility to
CC   lysis of tumour cells of various origins by hTRT CTL suggest that
CC   hTRT could serve as a universal cancer vaccine for humans.  A
CC   claimed universal vaccine for treating tumours of any origin
CC   comprises at least 1 hTRT peptide.  The peptide is 7-15 amino
CC   acid residues in length and may be modified to enhance binding to
CC   the major histocompatibility complex.  Also claimed is a method for
```

```
CC   inducing and enhancing a CTL response against cancer cells, involving
CC   harvesting blood leucocytes, pulsing with hTRT, and contacting
CC   cancer cells with the pulsed leucocytes.  A method for targeting
CC   CTL to tumour cells is also claimed, and involves administering a
CC   hTRT peptide to a mammal, especially a cancer patient.
XX
SQ   Sequence   9 AA;

  Query Match             88.9%;  Score 40;  DB 22;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADLNL 9
              |||||:|||
Db          1 YLSGANLNL 9


RESULT 15
AAE05123
ID   AAE05123 standard; peptide; 9 AA.
XX
AC   AAE05123;
XX
DT   18-SEP-2001  (first entry)
XX
DE   Carcinoembryonic antigen (CEA) peptide, CAP-1.
XX
KW   Tumour-associated antigen; TAA; cytostatic; vaccine; gene therapy;
KW   immune response; tetanus toxoid; TT; diphtheria toxoid; DT; prophylactic;
KW   cancer; therapeutic; carcinoembryonic antigen; CEA.
XX
OS   Unidentified.
XX
PN   WO200149317-A2.
XX
PD   12-JUL-2001.
XX
PF   05-JAN-2001; 2001WO-CA00005.
XX
PR   05-JAN-2000; 2000US-0174587.
XX
PA   (AVET ) AVENTIS PASTEUR LTD.
XX
PI   Emtage P,  Barber BH,  Sambhara S,  Sia CDY;
XX
DR   WPI; 2001-441790/47.
XX
PT   Enhancing immune response to antigen such as tumor antigen for treating
PT   cancer in an animal involves administering an inducing agent to the
PT   animal followed by administering inducing agent-antigen mixture  -
XX
PS   Example 2; Page 31; 62pp; English.
XX
CC   The invention relates to a method of enhancing an immune response against
CC   tumour-associated antigens (TAAs), such as GP100 and carcinoembryonic
CC   antigen (CEA) in an animal. The method involves priming of the animal
CC   with an inducing agent such as tetanus toxoid (TT) or diphtheria toxoid
CC   (DT), subsequently followed by administration of an inducing agent-
CC   antigen mixture. The method provides the enhancement or augmentation of
CC   the immune response to the antigen and/or improves a vaccination protocol
CC   by allowing use of less antigen. The immunisation of the animal with
CC   tumour-associated antigen is useful for the prophylactic or therapeutic
CC   treatment of cancer. The present sequence is carcinoembryonic antigen
CC   (CEA) peptide fragment related to the invention.
XX
SQ   Sequence   9 AA;

  Query Match             88.9%;  Score 40;  DB 22;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADLNL 9
              |||||:|||
Db          1 YLSGANLNL 9

Search completed: January 12, 2004, 14:28:23
Job time : 33.25 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         January 12, 2004, 14:21:08 ; Search time 32.25 Seconds
                                          (without alignments)
                                          44.296 Million cell updates/sec

Title:          US-09-529-121A-4
Perfect score:  45
Sequence:       1 YLSGANINL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters:      179625

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_19Jun03:*
                1:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
                2:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
                3:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
                4:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
                5:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
                6:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
                7:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
                8:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
                9:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
                10:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
                11:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
                12:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
                13:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                14:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
                15:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
                16:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                17:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
                18:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                19:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
                20:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                21:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
                22:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
                23:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
                24:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                       %
Result                Query
  No.     Score  Match Length DB  ID                       Description
-------------------------------------------------------------------------
    1       45   100.0      9  20  AAY09528                 Carcinoembryonic a
    2       43    95.6      9  18  AAW39723                 Human carcina-embr
    3       43    95.6      9  19  AAW77134                 CEA synthetic pept
    4       43    95.6      9  19  AAW70045                 CEA derived HLA-A2
    5       43    95.6      9  20  AAY47655                 Immunogenic peptid
    6       43    95.6      9  20  AAY09525                 Carcinoembryonic a
    7       43    95.6      9  21  AAB13749                 Peptide fragment #
    8       43    95.6      9  22  AAB82776                 Carcinoembryonic a
    9       43    95.6      9  22  AAE05123                 Carcinoembryonic a
   10       43    95.6      9  22  AAE02673                 Human CEA epitopic
   11       43    95.6      9  22  AAE00463                 Human tumour CEA e
   12       43    95.6      9  23  AAE26805                 Human HLA-A2.1 res
   13       43    95.6      9  23  ABG79073                 Human CEA class I
   14       43    95.6      9  23  AAU95893                 Immunogenic peptid
   15       43    95.6      9  23  AAE19088                 HLA-A24 restricted
   16       40    88.9      9  20  AAY09527                 Carcinoembryonic a
   17       40    88.9      9  21  AAY54173                 HLA binding peptid
   18       40    88.9      9  22  AAU26560                 Human Leukocyte An
   19       40    88.9      9  22  AAB99681                 HLA A2 binding CTL
   20       38    84.4      9  20  AAY09526                 Carcinoembryonic a
   21       38    84.4      9  21  AAB13750                 Peptide fragment #
   22       38    84.4      9  22  AAE05124                 Modified carcinoem
   23       38    84.4      9  22  AAB97818                 Carcinoembryonic a
   24       38    84.4      9  23  AAB47917                 Modified CEA epito
   25       38    84.4      9  23  AAE19089                 HLA-A24 restricted
   26       37    82.2      9  22  AAB75854                 Tumour associated
   27       36    80.0      9  17  AAW00680                 Peptide comprising
   28       34    75.6      9  20  AAY09529                 Carcinoembryonic a
   29       31    68.9      9  23  AAU82064                 CEA antigenic pept
   30       27    60.0      9  17  AAW00690                 NCA analogue of re
   31       26    57.8      9  22  AAU23920                 Human MHC class I
   32       26    57.8      9  22  AAU24043                 Human MHC class I
   33       26    57.8      9  22  AAU24335                 Human MHC class I
   34       26    57.8      9  24  ABJ19952                 MHC binding peptid
   35       23    51.1      9  19  AAW54298                 Human cytohesin-1
   36       23    51.1      9  21  AAB26339                 Human CASB618 prot
   37       23    51.1      9  23  AAM49862                 Human D40 associat
   38       23    51.1      9  24  ABP71902                 R. erythropolis AN
   39       23    51.1      9  24  ABP71908                 D. radiodurans crt
   40       22    48.9      7  21  AAB27207                 Reporter gene cons
   41       22    48.9      7  22  AAU09600                 Novel antiarrhythm
   42       22    48.9      8  21  AAB28155                 Linker between thy
   43       22    48.9      8  22  ABP18160                 HIV B58 super moti
   44       22    48.9      9  18  AAW19861                 Fragment of enzyme
   45       22    48.9      9  21  AAB26327                 Human CASB618 prot


                                   ALIGNMENTS


RESULT 1
AAY09528
ID   AAY09528 standard; peptide; 9 AA.
XX
AC   AAY09528;
XX
DT   20-JUL-1999  (first entry)
XX
DE   Carcinoembryonic antigen peptide agonist SEQ ID NO:4.
XX
KW   Carcinoembryonic antigen; CEA; human; agonist; antagonist;
KW   immune response; carcinoma; gastrointestinal; breast; pancreatic;
KW   bladder; ovarian; lung; prostatic; T cell proliferation; cancer;
KW   adoptive transfer therapy; autoimmune reaction; immunotherapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9919478-A1.
XX
PD   22-APR-1999.
XX
PF   22-SEP-1998;   98WO-US19794.
XX
PR   10-OCT-1997;   97US-0061589.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Barzaga E,  Schlom J,  Zaremba S;
XX
DR   WPI; 1999-326544/27.
XX
```

```
PT   Peptide agonists and antagonists of carcinoembryonal antigen
XX
PS   Claim 5; Page 53; 72pp; English.
XX
CC   The present invention describes peptides (A) that comprise agonists (Ia)
CC   or antagonists (Ib) of human carcinoembryonal antigen (CEA). (Ia) are
CC   used in vaccines to kill or inhibit carcinoma cells that express CEA or
CC   its epitopes, particularly for treating gastrointestinal, breast,
CC   pancreatic, bladder, ovarian, lung or prostatic carcinoma. They can also
CC   be used to proliferate T cells, e.g. from vaccinated subjects, for use
CC   in adoptive transfer therapy. (Ib) are used to inhibit CEA-specific
CC   immune responses, e.g. in vaccinated subjects, to prevent an autoimmune
CC   reaction to cancer immunotherapy (i.e. to prevent attack on normal but
CC   CEA-expressing cells). (Ia) are more active than native sequence (I) and
CC   generate a highly specific and systemic anti-CEA response. Cytotoxic T
CC   cells generated recognize both (Ia) and native CEA epitopes. The present
CC   sequence represents a specifically claimed example of (Ia).
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 45;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGANINL 9
              |||||||||
Db          1 YLSGANINL 9


RESULT 2
AAW39723
ID   AAW39723 standard; peptide; 9 AA.
XX
AC   AAW39723;
XX
DT   11-JUN-1998  (first entry)
XX
DE   Human carcina-embryonic antigen (CEA) peptide (pos. 571-579).
XX
KW   T cell epitope; immune response; human leukocyte antigen; HLA Class I;
KW   vaccine; immunogenic; major histocompatablilty complex; MHC; B cell;
KW   disease; anti-tumour; anti-viral.
XX
OS   Homo sapiens.
XX
PN   WO9741440-A1.
XX
PD   06-NOV-1997.
XX
PF   28-APR-1997;   97WO-NL00229.
XX
PR   23-DEC-1996;   96EP-0203670.
PR   26-APR-1996;   96EP-0201145.
XX
PA   (UYLE-) RIJKSUNIV LEIDEN.
PA   (SCIS-) SCI SEED CAPITAL INVESTMENTS BV.
XX
PI   Kast WM,  Melief CJM,  Offringa R,  Toes REM,  Van Der Burg SH;
XX
DR   WPI; 1997-549891/50.
XX
PT   Method of selecting T cell peptide epitope(s) - by measuring the
PT   stability of HLA class I-peptide complexes on intact B cells
XX
PS   Example 3; Page 85; 109pp; English.
XX
CC   Peptides AAW39430-W39734 are used in a novel method for the selection of
CC   immunogenic T-cell peptide epitopes present in polypeptide antigens. The
CC   method involves the identification of peptide sequences capable of
CC   binding to an HLA (human leukocyte antigen) class I molecule and
CC   measuring the binding of this epitope peptide to the HLA class I
CC   peptide. The stability of binding of the peptide and MHC (major
CC   histocompatablilty complex) class I molecule is measured on intact human
CC   B cells carrying the MHC molecule at their cell surfaces. The method can
CC   be used to select peptide epitopes for generating vaccines against a
CC   disease associated with the polypeptide, e.g. cancers or AIDS. The
CC   peptide epitopes are especially T-cell peptide epitopes with strong
CC   anti-tumour and anti-viral immune responses. Peptide AAW39723 is derived
CC   from the human carcino-embryonic antigen (CEA) and has the ability to
CC   bind to the human MHC Class I allele HLA-A2.1.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 18;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGANINL 9
              ||||||:||
Db          1 YLSGANLNL 9


RESULT 3
AAW77134
ID   AAW77134 standard; peptide; 9 AA.
XX
AC   AAW77134;
XX
DT   16-NOV-1998  (first entry)
XX
DE   CEA synthetic peptide epitope 1.
XX
KW   Tyrosinase; tyrosinase cytotoxic lymphocyte response;
KW   cytotoxic T lymphocyte; cysteine-depleted; melanoma.
XX
OS   Synthetic.
XX
PN   WO9833810-A2.
XX
PD   06-AUG-1998.
XX
PF   29-JAN-1998;   98WO-US01592.
XX
PR   30-JAN-1997;   97US-0037781.
XX
PA   (UYVI-) UNIV VIRGINIA PATENT FOUND.
XX
PI   Engelhard VH,  Hunt DF,  Kittlesen D,  Slingluff CL;
XX
DR   WPI; 1998-437388/37.
XX
PT   Disease specific immunogen  - comprises disease specific cytotoxic T
PT   lymphocyte epitope used to elicit melanoma specific CTL response
XX
PS   Disclosure; Page 27; 93pp; English.
XX
CC   The peptide epitope AAW77119-W77138 were created for human
CC   tumour-specific cytotoxic T lymphocyte response.  These peptides are are
CC   cysteine- depleted mutants of a native disease-specific CTL epitope.  The
CC   cysteine- depleted CTL epitopes elicit a stronger or more specific CTL
CC   response than the native epitope.  The epitopes can be used in a
CC   disease-specific immunogen to protect a mammal against disease in
CC   particular melanomas. The peptides may also be used to screen a sample
CC   for the presence of an antigen with the same epitope, or with a different
CC   cross-reactive epitope.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 19;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGANINL 9
              ||||||:||
Db          1 YLSGANLNL 9
```

```
RESULT 4
AAW70045
ID   AAW70045 standard; peptide; 9 AA.
XX
AC   AAW70045;
XX
DT   22-OCT-1998  (first entry)
XX
DE   CEA derived HLA-A2.1 binding peptide 2 (residues 605-613).
XX
KW   Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW   human leukocyte antigen; HLA; tumour associated antigen; cancer;
KW   antigen presenting cell; APC; immunogenic peptide; immune disorder;
KW   viral infection; AIDS; hepatitis; bacterial infection; malaria; CEA;
KW   fungal infection; tuberculosis; melanoma; carcinoembryonic antigen.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9833888-A1.
XX
PD   06-AUG-1998.
XX
PF   30-JAN-1998;   98WO-US01959.
XX
PR   31-JAN-1997;   97US-0036696.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Celis E,  Sette A,  Sidney J,  Southwood S,  Tsai V;
XX
DR   WPI; 1998-437445/37.
XX
PT   Production of antigen-specific cytotoxic T cells - by incubating
PT   immunogenic peptide(s) from antigen that binds class I major
PT   histocompatibility complex molecules with pre-treated antigen
PT   presenting cells
XX
PS   Example 6; Page 75; 104pp; English.
XX
CC   Sequences shown in AAW70044 to AAW70052 represent peptides derived from
CC   carcinoembryonic antigen (CEA). The peptides can bind to a human
CC   leukocyte antigen (HLA), HLA-A2.1 and are used to exemplify the method
CC   of invention of producing antigen-specific cytotoxic T cells (CTLs) in
CC   vitro. The method comprises contacting immunogenic peptides from an
CC   antigen that binds class I major histocompatibility complex (MHC)
CC   molecules with antigen presenting cells (APCs) pretreated with
CC   pretreatment growth factors, and incubating the APCs with purified CD8
CC   cells in the presence of at least 2 incubation growth factors, thereby
CC   producing antigen-specific CTLs. A method for specifically killing
CC   target cells in a human patient is also provided which comprises
CC   obtaining a fluid sample containing CTLs from a patient, contacting the
CC   cytotoxic T cells with APCs pretreated with pre-treatment growth
CC   factors, where the APCs comprise class I MHC molecules. The pretreated
CC   APCs are incubated with the cytotoxic growth factors, thereby producing
CC   activated CTLs which are contacted with a carrier to form a composition.
CC   The composition can then be administered to the patient. The activated
CC   CTLs can be used for treating cancers, immune disorders, viral
CC   infections, AIDS, hepatitis, bacterial infection, fungal infection,
CC   malaria or tuberculosis.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 19;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGANINL 9
              ||||||:||
Db          1 YLSGANLNL 9
```

```
RESULT 5
AAY47655
ID   AAY47655 standard; Peptide; 9 AA.
XX
AC   AAY47655;
XX
DT   01-DEC-1999  (first entry)
XX
DE   Immunogenic peptide having a human leukocyte antigen binding motif #2266.
XX
KW   Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW   immune response; T cell activation; major histocompatibility complex;
KW   cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW   prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW   vaccine; immunisation.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9945954-A1.
XX
PD   16-SEP-1999.
XX
PF   13-MAR-1998;   98WO-US05039.
XX
PR   13-MAR-1998;   98WO-US05039.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Kubo RT,  Sidney J,  Celis E,  Grey HM,  Southwood S;
XX
DR   WPI; 1999-551214/46.
XX
PT   New immunogenic peptides with HLA binding motif, useful in treatment
PT   and diagnosis of cancers and viral diseases  -
XX
PS   Claim 1; Page 118; 150pp; English.
XX
CC   AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
CC   having a human major histocompatibility complex (MHC) Class I (also
CC   known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC   peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC   HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC   response against the antigen from which the peptide is derived.
CC   Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC   normally induced by an antigen in the form of a peptide fragment bound
CC   to a HLA molecule, rather than the intact foreign antigen itself, and
CC   are particularly important in tumour rejection and in fighting viral
CC   infections. The peptides are therefore useful therapeutically to treat
CC   or prevent viral infections and cancers in mammals (especially humans)
CC   e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC   They can be administered as vaccines to elicit an immune response in
CC   individuals susceptible or otherwise at risk of viral infection or
CC   cancer, or used to treat chronic or acute conditions. They are also
CC   useful diagnostically, and can be used to induce a cytotoxic T cell
CC   response, by contacting a cytotoxic T cell with the peptide e.g. to
CC   produce CTLs ex vivo for infusion back into a patient. The
CC   polynucleotides encoding the immunogenic peptides are also useful
CC   therapeutically and for immunisation as above.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 20;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGANINL 9
              ||||||:||
Db          1 YLSGANLNL 9


RESULT 6
AAY09525
```

```
ID   AAY09525 standard; peptide; 9 AA.
XX
AC   AAY09525;
XX
DT   20-JUL-1999  (first entry)
XX
DE   Carcinoembryonic antigen peptide agonist CAP-1.
XX
KW   Carcinoembryonic antigen; CEA; human; agonist; antagonist;
KW   immune response; carcinoma; gastrointestinal; breast; pancreatic;
KW   bladder; ovarian; lung; prostatic; T cell proliferation; cancer;
KW   adoptive transfer therapy; autoimmune reaction; immunotherapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9919478-A1.
XX
PD   22-APR-1999.
XX
PF   22-SEP-1998;   98WO-US19794.
XX
PR   10-OCT-1997;   97US-0061589.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Barzaga E,  Schlom J,  Zaremba S;
XX
DR   WPI; 1999-326544/27.
XX
PT   Peptide agonists and antagonists of carcinoembryonal antigen
XX
PS   Claim 1; Page 53; 72pp; English.
XX
CC   The present invention describes peptides (A) that comprise agonists (Ia)
CC   or antagonists (Ib) of human carcinoembryonal antigen (CEA). (Ia) are
CC   used in vaccines to kill or inhibit carcinoma cells that express CEA or
CC   its epitopes, particularly for treating gastrointestinal, breast,
CC   pancreatic, bladder, ovarian, lung or prostatic carcinoma. They can also
CC   be used to proliferate T cells, e.g. from vaccinated subjects, for use
CC   in adoptive transfer therapy. (Ib) are used to inhibit CEA-specific
CC   immune responses, e.g. in vaccinated subjects, to prevent an autoimmune
CC   reaction to cancer immunotherapy (i.e. to prevent attack on normal but
CC   CEA-expressing cells). (Ia) are more active than native sequence (I) and
CC   generate a highly specific and systemic anti-CEA response. Cytotoxic T
CC   cells generated recognize both (Ia) and native CEA epitopes. The present
CC   sequence represents a specifically claimed example of (Ia).
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 20;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGANINL 9
              ||||||:||
Db          1 YLSGANLNL 9

RESULT 7
AAB13749
ID   AAB13749 standard; peptide; 9 AA.
XX
AC   AAB13749;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Peptide fragment # 1 from human CEA.
XX
KW   Human; T-cell; immune response; antigen; epitope; B7 family molecule;
KW   Leukocyte function-associated antigen-3; LFA-3;
KW   Intercellular adhesion molecule-1; ICAM-1; vaccine; immunotherapy;
KW   colon polyp; Crohn's disease; ulcerative colitis; breast lesion;
```

```
KW   tumour; CEA.
XX
OS   Homo sapiens.
XX
PN   WO200034494-A1.
XX
PD   15-JUN-2000.
XX
PF   12-NOV-1999;   99WO-US26866.
XX
PR   09-DEC-1998;   98US-0111582.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (THER-) THERION BIOLOGICS CORP.
XX
PI   Schlom J,  Hodge J,  Panicali D;
XX
DR   WPI; 2000-431307/37.
XX
PT   Novel recombinant vector useful as immunogens and vaccines for
PT   stimulating and enhancing immunological responses to target cells and
PT   antigens expresses multiple co-stimulatory molecules such as B7-1,
PT   LFA-3, ICAM-1  -
XX
PS   Claim 18; Page 35; 188pp; English.
XX
CC   Costimulatory molecules have important roles in T-cell activation and
CC   therefore the immune response. The present invention relates to
CC   recombinant vectors which comprise of foreign nucleic acid sequences
CC   encoding at least three costimulatory molecules: a B7 family molecule,
CC   Leukocyte function-associated antigen-3 (LFA-3, human CD58) and
CC   Intercellular adhesion molecule-1 (ICAM-1, CD54) and optionally a foreign
CC   gene encoding a target antigen or immunological epitope. The present
CC   sequence is one such target antigen used in the present invention. The
CC   present sequence is a tumour-associated antigen. The vector of the
CC   present invention would be useful for providing an enhanced immune
CC   response to the present target antigen. The vector of the present
CC   invention may therefore be useful in immunotherapy for treating or
CC   preventing diseases caused by viruses, bacteria, protozoans, parasites,
CC   premalignant cells and tumour cells. The recombinant vector can be used
CC   to treat or prevent preneoplastic or hyperplastic states such as colon
CC   polyps, Crohn's disease, ulcerative colitis and breast lesions.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 21;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGANINL 9
              ||||||:||
Db          1 YLSGANLNL 9

RESULT 8
AAB82776
ID   AAB82776 standard; Protein; 9 AA.
XX
AC   AAB82776;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Carcinoembryonic antigen peptide.
XX
KW   Telomerase reverse transcriptase; hTRT; human;
KW   cytotoxic T lymphocyte; major histocompatibility complex; cancer;
KW   tumour; human leucocyte antigen; HLA-A2.1; vaccine;
KW   carcinoembryonic antigen.
XX
OS   Homo sapiens.
XX
PN   WO200160391-A1.
XX
```

```
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05143.
XX
PR   15-FEB-2000; 2000US-0182685.
PR   15-FEB-2001; 2001US-0182685.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Zanetti M;
XX
DR   WPI; 2001-536552/59.
XX
PT   Vaccine for initiating and enhancing a cytotoxic T lymphocyte response,
PT   for treating cancers or tumours or for inducing immune response against
PT   tumours, comprises a telomerase reverse transcriptase peptide    -
XX
PS   Example 1; Page 12; 52pp; English.
XX
CC   The present sequence is that of a carcinoembryonic antigen peptide
CC   comprising amino acid residues 571-579.  The peptide was used as a
CC   reference peptide in comparison with human telomerase reverse
CC   transcriptase (hTRT) HLA-A2.1+ restricted peptide p540 (see
CC   AAB82772) in a HLA-A2.1 binding/stabilisation assay.  The induction
CC   of CTL responses in vitro and in vivo, and the susceptibility to
CC   lysis of tumour cells of various origins by hTRT CTL suggest that
CC   hTRT could serve as a universal cancer vaccine for humans.  A
CC   claimed universal vaccine for treating tumours of any origin
CC   comprises at least 1 hTRT peptide.  The peptide is 7-15 amino
CC   acid residues in length and may be modified to enhance binding to
CC   the major histocompatibility complex.  Also claimed is a method for
CC   inducing and enhancing a CTL response against cancer cells, involving
CC   harvesting blood leucocytes, pulsing with hTRT, and contacting
CC   cancer cells with the pulsed leucocytes.  A method for targeting
CC   CTL to tumour cells is also claimed, and involves administering a
CC   hTRT peptide to a mammal, especially a cancer patient.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 22;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGANINL 9
              ||||||:||
Db          1 YLSGANLNL 9


RESULT 9
AAE05123
ID   AAE05123 standard; peptide; 9 AA.
XX
AC   AAE05123;
XX
DT   18-SEP-2001  (first entry)
XX
DE   Carcinoembryonic antigen (CEA) peptide, CAP-1.
XX
KW   Tumour-associated antigen; TAA; cytostatic; vaccine; gene therapy;
KW   immune response; tetanus toxoid; TT; diphtheria toxoid; DT; prophylactic;
KW   cancer; therapeutic; carcinoembryonic antigen; CEA.
XX
OS   Unidentified.
XX
PN   WO200149317-A2.
XX
PD   12-JUL-2001.
XX
PF   05-JAN-2001; 2001WO-CA00005.
XX
PR   05-JAN-2000; 2000US-0174587.
XX
PA   (AVET ) AVENTIS PASTEUR LTD.
```

```
XX
PI   Emtage P,  Barber BH,  Sambhara S,  Sia CDY;
XX
DR   WPI; 2001-441790/47.
XX
PT   Enhancing immune response to antigen such as tumor antigen for treating
PT   cancer in an animal involves administering an inducing agent to the
PT   animal followed by administering inducing agent-antigen mixture  -
XX
PS   Example 2; Page 31; 62pp; English.
XX
CC   The invention relates to a method of enhancing an immune response against
CC   tumour-associated antigens (TAAs), such as GP100 and carcinoembryonic
CC   antigen (CEA) in an animal. The method involves priming of the animal
CC   with an inducing agent such as tetanus toxoid (TT) or diphtheria toxoid
CC   (DT), subsequently followed by administration of an inducing agent-
CC   antigen mixture. The method provides the enhancement or augmentation of
CC   the immune response to the antigen and/or improves a vaccination protocol
CC   by allowing use of less antigen. The immunisation of the animal with
CC   tumour-associated antigen is useful for the prophylactic or therapeutic
CC   treatment of cancer. The present sequence is carcinoembryonic antigen
CC   (CEA) peptide fragment related to the invention.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 22;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGANINL 9
              ||||||:||
Db          1 YLSGANLNL 9


RESULT 10
AAE02673
ID   AAE02673 standard; peptide; 9 AA.
XX
AC   AAE02673;
XX
DT   06-AUG-2001  (first entry)
XX
DE   Human CEA epitopic peptide.
XX
KW   Human; cytostatic; antibacterial; antifungal; gene therapy; vaccine;
KW   antiviral; tumour; epitope; glycoprotein; hepatitis B virus; HBV;
KW   immune response; CTL; cytotoxic T lymphocyte; CEA; HLA;
KW   human leucocyte antigen.
XX
OS   Homo sapiens.
XX
PN   WO200127291-A1.
XX
PD   19-APR-2001.
XX
PF   29-SEP-2000; 2000WO-EP09902.
XX
PR   12-OCT-1999;   99US-0158356.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Firat H,  Lemonnier F,  Langlade-demoyen P;
XX
DR   WPI; 2001-282038/29.
XX
PT   New polynucleotide comprising at least one viral, fungal, bacterial, or
PT   tumor epitope of an antigen, capable of inducing a cellular response -
XX
PS   Example 1; Page 23; 70pp; English.
XX
CC   The invention relates to polynucleotide containing at least a part of
CC   the coding sequence of the middle glycoprotein of hepatitis B virus
CC   (HBV) in which is inserted a DNA sequence coding for an epitope
```

```
CC   comprising at least one viral, fungal, bacterial, or tumour epitope of
CC   an antigen, capable of inducing a cellular response. Nucleic acids and
CC   compositions of the invention are useful for inducing in vivo a CTL
CC   (cytotoxic T lymphocyte) response against several epitopes of one
CC   or more, bacterial, viral, fungal, or tumour antigens. A composition
CC   of the invention produces an immune response against HIV antigen and
CC   are used in the production of vaccines. The polynucleotides of the
CC   invention are also used in gene therapy. The present sequence is
CC   human CEA epitopic peptide. This peptide elicits strong
CC   HLA (human leucocyte antigen)-A2.1-restricted CTL response in mice.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 22;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGANINL 9
              ||||||:||
Db          1 YLSGANLNL 9


RESULT 11
AAE00463
ID   AAE00463 standard; peptide; 9 AA.
XX
AC   AAE00463;
XX
DT   19-JUN-2001  (first entry)
XX
DE   Human tumour CEA epitopic peptide.
XX
KW   Human; tumour epitope; cytostatic; immunostimulant; gene therapy;
KW   middle glycoprotein; Hepatitis B virus; HBV; cytotoxic response;
KW   immune response; cytotoxic T lymphocyte; CTL; CEA; HLA;
KW   human leucocyte antigen.
XX
OS   Homo sapiens.
XX
PN   WO200123577-A2.
XX
PD   05-APR-2001.
XX
PF   29-SEP-2000; 2000WO-EP09900.
XX
PR   30-SEP-1999;   99US-0156945.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Firat H,  Lemonnier F,  Langlade-demoyen P,  Michel M,  Suhrbier AA;
XX
DR   WPI; 2001-266164/27.
XX
PT   Novel polynucleotide having DNA sequence encoding tumor antigen epitope
PT   inserted in part of coding sequence of middle glycoprotein of hepatitis
PT   B virus, used to induce immune response against tumor-specific antigen
PT   -
XX
PS   Example 1; Page 13; 36pp; English.
XX
CC   The present invention relates to an isolated or purified polynucleotide
CC   containing a DNA sequence coding for at least one tumour epitope of a
CC   tumour antigen inserted into part of the coding sequence of the middle
CC   glycoprotein of the Hepatitis B virus (HBV). The polynucleotide is
CC   useful for optionally evaluating cytotoxic responses in the individual's
CC   lymphocyte population. It induces an immune response against at least
CC   one tumour specific antigen or tissue specific antigen. The vector
CC   comprising the polynucleotide induces in vivo, cellular and/or humoral
CC   immune response. The composition comprising the polynucleotide induces
CC   in vivo, cytotoxic T lymphocyte (CTL) against one or more antigens or
CC   epitopes present on the hybrid protein. The polynucleotide is also
CC   useful in gene therapy.
CC   The present sequence is a human tumour CEA epitopic peptide. This
```

```
CC   peptide elicits strong HLA (human leucocyte antigen)-A2.1-restricted
CC   CTL response in mice.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 22;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGANINL 9
              ||||||:||
Db          1 YLSGANLNL 9


RESULT 12
AAE26805
ID   AAE26805 standard; peptide; 9 AA.
XX
AC   AAE26805;
XX
DT   13-DEC-2002  (first entry)
XX
DE   Human HLA-A2.1 restricted CEA (carcinoembryonic Ag) peptide epitope.
XX
KW   Human; cancer; breast cancer; ovarian cancer; melanoma; cell therapy;
KW   epitope; human leucocyte antigen; HLA-A2.1.
XX
OS   Homo sapiens.
XX
PN   WO200265992-A2.
XX
PD   29-AUG-2002.
XX
PF   19-FEB-2002; 2002WO-US05748.
XX
PR   20-FEB-2001; 2001US-270252P.
XX
PA   (ORTH ) ORTHO-MCNEIL PHARM INC.
XX
PI   Degraw J,  Moriarty A,  Leturcq DJ,  Jackson MR,  Peterson PA;
PI   Heiskala M;
XX
DR   WPI; 2002-667033/71.
XX
PT   Treating a subject with cancer comprises combining the CD+8 cells,
PT   which are stimulated with non-naturally occurring antigen-presenting
PT   cell line, with adherent blood monocytes and inoculating the subject
PT   with CD8+ suspension  -
XX
PS   Example 2; Page 93; 99pp; English.
XX
CC   The invention relates to a method of treating a subject with cancer. The
CC   method involves combining the CD+8 cells, which are stimulated with non
CC   naturally occurring antigen-presenting cell (nnAPC) line, with adherent
CC   blood monocytes and inoculating the subject with CD8+ suspension. The
CC   method is useful for treating cancer e.g. ovarian cancer, breast cancer
CC   and melanoma etc. It is also useful in cell therapy. The present sequence
CC   is human leukocyte antigen A2 (HLA-A2).1 restricted peptide epitope used
CC   to treat breast and ovarian cancer.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 23;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGANINL 9
              ||||||:||
Db          1 YLSGANLNL 9


RESULT 13
ABG79073
```

```
ID   ABG79073 standard; Peptide; 9 AA.
XX
AC   ABG79073;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Human CEA class I HLA widely expressed antigen peptide #1.
XX
KW   Cell penetrating peptide; cancer; tumour; melanoma; thymoma; antigen;
KW   lymphoma; sarcoma; lung cancer; non-Hodgkin's lymphoma; leukaemia;
KW   Hodgkin's lymphoma; uterine cancer; cervical cancer; bladder cancer;
KW   kidney cancer; adenocarcinoma; breast cancer; prostate cancer;
KW   ovarian cancer; pancreatic cancer; epitope; vaccine; dendritic cell;
KW   tumour infiltrating lymphocyte; TIL; human leukocyte antigen; HLA;
KW   cytostatic; human.
XX
OS   Homo sapiens.
XX
PN   WO200264057-A2.
XX
PD   22-AUG-2002.
XX
PF   15-FEB-2002; 2002WO-US05212.
XX
PR   15-FEB-2001; 2001US-268687P.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI   Wang R;
XX
DR   WPI; 2002-627577/67.
XX
PT   Novel composition for treating a disease in an animal, comprises an
PT   immune effector cell and cell penetrating peptide associated with an
PT   antigen or antibody -
XX
PS   Disclosure; Page 17; 61pp; English.
XX
CC   The invention relates to a composition (I) comprising an immune effector
CC   cell and a cell penetrating peptide (CPP) associated with an antigen or
CC   antibody. Also included are (1) a vaccine comprising (I), CPP
CC   associated with an antigen, and a pharmaceutically acceptable carrier
CC   and (2) preparing a composition for a disease, by providing (I)
CC   and CPP associated with an antigen for disease, and introducing the
CC   antigen-associated CPP to (I), where antigen enters into the cell.
CC   The antigens are, for example, tumour antigen derived epitopes
CC   recognised by tumour infiltrating lymphocytes (TIL) of HLA (human
CC   leukocyte antigen) class I or II. The composition is useful for enhancing
CC   immunity in an animal to a disease, by administering a mature dendritic
CC   cell comprising CPP associated with an antigen to disease, to the animal,
CC   such that following the administration, animal is protected from disease,
CC   where the animal comprises both CD4+ and CD8+ T cells. It is also useful
CC   for treating a disease (e.g. cancer, tumour, melanoma, thymoma,
CC   lymphoma, sarcoma, lung cancer, non-Hodgkin's lymphoma, leukaemia,
CC   Hodgkin's lymphoma, uterine cancer, cervical cancer, bladder cancer,
CC   kidney cancer, adenocarcinoma, breast cancer, prostate cancer,
CC   ovarian cancer and pancreatic cancer). The animal is further subjected to
CC   a cancer treatment including surgery, radiation, chemotherapy or gene
CC   therapy. The administration of (I), preferably dendritic cell is prior
CC   to, subsequent to or concurrent with, the cancer treatment. The present
CC   sequence is a tumour antigen derived epitope for inclusion in the
CC   composition of the invention.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 23;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGANINL 9
              ||||||:||
Db          1 YLSGANLNL 9
```

```
RESULT 14
AAU95893
ID   AAU95893 standard; Peptide; 9 AA.
XX
AC   AAU95893;
XX
DT   02-JUL-2002  (first entry)
XX
DE   Immunogenic peptide with (HLA)-A2.1 binding site #106.
XX
KW   HLA-A2.1 binding peptide; cytostatic; virucide; anti-HIV; hepatotropic;
KW   human immunodeficiency virus; antiinflammatory; antibacterial; vaccine;
KW   protozoacide; immunosuppressant; immunogenic peptide; T cell activation;
KW   human leucocyte antigen binding site; cytotoxic T cell response;
KW   viral infection; hepatitis; Epstein-Barr virus; papilloma virus;
KW   human immunodeficiency virus; HIV; Kaposi sarcoma; Lassa fever virus;
KW   cytomegalovirus; tumour; prostate cancer; renal carcinoma; lymphona;
KW   prostate-specific antigen; p53; carcino-embryonal antigen;
KW   melanoma antigen; Mycobacterium tuberculosis; protozoa;
KW   trypanosome surface antigen; condlyloma acuminatum.
XX
OS   Unidentified.
XX
PN   WO200220616-A1.
XX
PD   14-MAR-2002.
XX
PF   01-SEP-2000; 2000WO-US24102.
XX
PR   01-SEP-2000; 2000WO-US24102.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Grey HM,  Sette A,  Sidney J,  Southwood S;
XX
DR   WPI; 2002-351766/38.
XX
PT   Immunogenic peptide with human leucocyte antigen-A2.1 binding site,
PT   useful for treating e.g. viral infection or tumours  -
XX
PS   Claim 1; Page 27; 35pp; English.
XX
CC   The invention describes a composition comprising an immunogenic peptide
CC   having a human leucocyte antigen (HLA)-A2.1 binding site. The peptides
CC   bind specifically to HLA-A2.1, to cause T cell activation and thus a
CC   cytotoxic T cell response. The peptides and the nucleic acids that
CC   encodes them, are used, in vivo or ex vivo, for treatment of viral
CC   infections (hepatitis B or C; Epstein-Barr; human immune deficiency;
CC   Kaposi sarcoma; human papilloma; Lassa fever or cytomegaloviruses);
CC   tumours including prostate cancer, renal carcinoma and lymphona (where
CC   directed to prostate-specific antigen, p53, carcino-embryonal antigen,
CC   Her2/neu or melanoma antigens); infection by Mycobacterium tuberculosis
CC   or protozoa (directed to trypanosome surface antigen); and condlyloma
CC   acuminatum. The peptides are suitable for use in peptide-based
CC   vaccines. This sequence represents an immunogenic peptide with the
CC   human leucocyte antigen (HLA)-A2.1 binding site, described in the
CC   invention.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 23;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGANINL 9
              ||||||:||
Db          1 YLSGANLNL 9


RESULT 15
AAE19088
ID   AAE19088 standard; peptide; 9 AA.
```

```
XX
AC   AAE19088;
XX
DT   21-MAY-2002  (first entry)
XX
DE   HLA-A24 restricted target antigen CEA immunological epitope #2.
XX
KW   Human leukocyte antigen; HLA; pharmaceutical composition; target antigen;
KW   immunological epitope; replication-defective virus; RDV; immune response;
KW   chemotherapy; granulocyte-monocyte-colony stimulating factor; cytostatic;
KW   GM-CSF; MHC; major histocompatibility complex; tumour; head; pancreatic;
KW   neck; breast; prostate; colorectal; melanoma; myeloidysplastic syndrome;
KW   metastatic breast skin lesion; corticosteroid therapy; erythropoietin;
KW   cytopenia; neutropenia; vaccine; immunostimulant.
XX
OS   Homo sapiens.
XX
PN   WO200195919-A2.
XX
PD   20-DEC-2001.
XX
PF   15-JUN-2001; 2001WO-US19201.
XX
PR   15-JUN-2000; 2000US-211717P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (THER-) THERION BIOLOGICS CORP.
XX
PI   Schlom J,  Greiner JW,  Kass E,  Panicali D;
XX
DR   WPI; 2002-205852/26.
XX
PT   Composition for enhancing immune responses, particularly anti-tumor
PT   responses and treating neutropenia, cytopenia, comprises
PT   replication-defective virus encoding granulocyte-monocyte-colony
PT   stimulating factor  -
XX
PS   Claim 9; Page 15; 118pp; English.
XX
CC   The present invention relates to a pharmaceutical composition comprising
CC   a replication-defective virus (RDV) encoding granulocyte-monocyte-colony
CC   stimulating factor (GM-CSF). The invention is useful for enhancing cell-
CC   mediated or humoral immune response in an individual, by enhancing
CC   migration of APC expressing CD11c^+/I-Ab^+, major histocompatibility
CC   complex (MHC) class II, at an injection site, regional lymph node at a
CC   tumour site, APC proliferation or function, CD4^+T or CD8^+T cell
CC   activation, interleukin (IL)-2, interferon (IFN)-gamma or tumour necrosis
CC   factor (TNF)-alpha production or their combinations. The composition
CC   enhances an antigen-specific T-cell response in an individual to a target
CC   antigen or its immunological epitope and an anti-tumour response in an
CC   individual with a head tumour, neck, breast, pancreatic, prostate,
CC   colorectal or metastatic tumour or melanoma, or metastatic breast skin
CC   lesion. The invention is further useful for treating neutropenia
CC   resulting from chemotherapy, corticosteroid therapy, irradiation or an
CC   infection, by raising the neutrophil count to normal levels and for
CC   treating cytopenias in patients with myeloidysplastic syndrome in
CC   combination with erythropoietin, by increasing neutrophil count and
CC   erythroid precursors. The composition enhances immune response to
CC   vaccines such as DPT, Td, DtaP, Hib, DtaP-Hib, MMR, Hepatitis A,
CC   hepatitis B, lyme's disease, influenza, tetravalent meningococcal
CC   polysaccharide, pneumococcal polysaccharide, anthrax, cholera, plague,
CC   yellow fever and Bacillus Calmette-Guerin vaccine. The present sequence
CC   is human leukocyte antigen (HLA)-restricted target tumour antigen
CC   immunological epitope.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 23;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGANINL 9
                |||||||:||
Db          1 YLSGANLNL 9

Search completed: January 12, 2004, 14:28:24
Job time : 33.25 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          January 12, 2004, 14:21:08 ; Search time 32.25 Seconds
                                             (without alignments)
                                             44.296 Million cell updates/sec

Title:           US-09-529-121A-5
Perfect score:   48
Sequence:        1 YLSGACLNL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters:      179625

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        A_Geneseq_19Jun03:*
                 1:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
                 2:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
                 3:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
                 4:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
                 5:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
                 6:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
                 7:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
                 8:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
                 9:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
                 10:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
                 11:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
                 12:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
                 13:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                 14:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
                 15:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
                 16:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                 17:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
                 18:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                 19:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
                 20:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                 21:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
                 22:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
                 23:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
                 24:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                    SUMMARIES
                        %
Result                Query
  No.    Score   Match Length DB  ID                   Description
 -------------------------------------------------------------------------
    1      48    100.0      9  20  AAY09529             Carcinoembryonic a
    2      36     75.0      9  18  AAW39723             Human carcina-embr
    3      36     75.0      9  19  AAW77134             CEA synthetic pept
    4      36     75.0      9  19  AAW70045             CEA derived HLA-A2
    5      36     75.0      9  20  AAY47655             Immunogenic peptid
    6      36     75.0      9  20  AAY09525             Carcinoembryonic a
    7      36     75.0      9  20  AAY09526             Carcinoembryonic a
    8      36     75.0      9  21  AAB13749             Peptide fragment #
    9      36     75.0      9  21  AAB13750             Peptide fragment #
```

```
   10      36     75.0      9  22  AAB82776             Carcinoembryonic a
   11      36     75.0      9  22  AAE05123             Carcinoembryonic a
   12      36     75.0      9  22  AAE05124             Modified carcinoem
   13      36     75.0      9  22  AAB97818             Carcinoembryonic a
   14      36     75.0      9  22  AAE02673             Human CEA epitopic
   15      36     75.0      9  22  AAE00463             Human tumour CEA e
   16      36     75.0      9  23  AAE26805             Human HLA-A2.1 res
   17      36     75.0      9  23  ABG79073             Human CEA class I
   18      36     75.0      9  23  AAU95893             Immunogenic peptid
   19      36     75.0      9  23  AAB47917             Modified CEA epito
   20      36     75.0      9  23  AAE19088             HLA-A24 restricted
   21      36     75.0      9  23  AAE19089             HLA-A24 restricted
   22      34     70.8      9  20  AAY09527             Carcinoembryonic a
   23      34     70.8      9  20  AAY09528             Carcinoembryonic a
   24      33     68.8      9  21  AAY54173             HLA binding peptid
   25      33     68.8      9  22  AAU26560             Human Leukocyte An
   26      33     68.8      9  22  AAB99681             HLA A2 binding CTL
   27      30     62.5      9  22  AAB75854             Tumour associated
   28      29     60.4      9  17  AAW00680             Peptide comprising
   29      26     54.2      7  19  AAW76240             Bacterial periplas
   30      26     54.2      9  21  AAY82882             Teratocarcinoma-de
   31      26     54.2      9  23  AAO14704             Human cripto prote
   32      24     50.0      7  20  AAY49030             Membrane dipeptida
   33      24     50.0      7  20  AAY39584             CTLA-4 VLD CDR1 re
   34      24     50.0      9  19  AAW70043             MAGE 3 antigen der
   35      24     50.0      9  20  AAY47153             Immunogenic peptid
   36      24     50.0      9  20  AAY47480             Immunogenic peptid
   37      24     50.0      9  20  AAY47562             Immunogenic peptid
   38      24     50.0      9  20  AAY47563             Immunogenic peptid
   39      24     50.0      9  22  AAG62724             Amino acid sequenc
   40      24     50.0      9  22  AAG84682             MAGE3 crossbinding
   41      24     50.0      9  22  AAG84848             MAGE3 HLA-A2 super
   42      24     50.0      9  23  AAO14721             Human cripto prote
   43      24     50.0      9  23  AAU82064             CEA antigenic pept
   44      23     47.9      6  17  AAY01944             Fragment of the EG
   45      23     47.9      6  22  AAU01136             Unknown peptide #3


                                  ALIGNMENTS


RESULT 1
AAY09529
ID   AAY09529 standard; peptide; 9 AA.
XX
AC   AAY09529;
XX
DT   20-JUL-1999  (first entry)
XX
DE   Carcinoembryonic antigen peptide agonist SEQ ID NO:5.
XX
KW   Carcinoembryonic antigen; CEA; human; agonist; antagonist;
KW   immune response; carcinoma; gastrointestinal; breast; pancreatic;
KW   bladder; ovarian; lung; prostatic; T cell proliferation; cancer;
KW   adoptive transfer therapy; autoimmune reaction; immunotherapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9919478-A1.
XX
PD   22-APR-1999.
XX
PF   22-SEP-1998;   98WO-US19794.
XX
PR   10-OCT-1997;   97US-0061589.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Barzaga E,  Schlom J,  Zaremba S;
XX
DR   WPI; 1999-326544/27.
XX
```

```
PT   Peptide agonists and antagonists of carcinoembryonal antigen
XX
PS   Claim 5; Page 53; 72pp; English.
XX
CC   The present invention describes peptides (A) that comprise agonists (Ia)
CC   or antagonists (Ib) of human carcinoembryonal antigen (CEA). (Ia) are
CC   used in vaccines to kill or inhibit carcinoma cells that express CEA or
CC   its epitopes, particularly for treating gastrointestinal, breast,
CC   pancreatic, bladder, ovarian, lung or prostatic carcinoma. They can also
CC   be used to proliferate T cells, e.g. from vaccinated subjects, for use
CC   in adoptive transfer therapy. (Ib) are used to inhibit CEA-specific
CC   immune responses, e.g. in vaccinated subjects, to prevent an autoimmune
CC   reaction to cancer immunotherapy (i.e. to prevent attack on normal but
CC   CEA-expressing cells). (Ia) are more active than native sequence (I) and
CC   generate a highly specific and systemic anti-CEA response. Cytotoxic T
CC   cells generated recognize both (Ia) and native CEA epitopes. The present
CC   sequence represents a specifically claimed example of (Ia).
XX
SQ   Sequence   9 AA;

  Query Match              100.0%;  Score 48;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGACLNL 9
              |||||||||
Db          1 YLSGACLNL 9


RESULT 2
AAW39723
ID   AAW39723 standard; peptide; 9 AA.
XX
AC   AAW39723;
XX
DT   11-JUN-1998  (first entry)
XX
DE   Human carcina-embryonic antigen (CEA) peptide (pos. 571-579).
XX
KW   T cell epitope; immune response; human leukocyte antigen; HLA Class I;
KW   vaccine; immunogenic; major histocompatablilty complex; MHC; B cell;
KW   disease; anti-tumour; anti-viral.
XX
OS   Homo sapiens.
XX
PN   WO9741440-A1.
XX
PD   06-NOV-1997.
XX
PF   28-APR-1997;   97WO-NL00229.
XX
PR   23-DEC-1996;   96EP-0203670.
PR   26-APR-1996;   96EP-0201145.
XX
PA   (UYLE-) RIJKSUNIV LEIDEN.
PA   (SCIS-) SCI SEED CAPITAL INVESTMENTS BV.
XX
PI   Kast WM,  Melief CJM,  Offringa R,  Toes REM,  Van Der Burg SH;
XX
DR   WPI; 1997-549891/50.
XX
PT   Method of selecting T cell peptide epitope(s) - by measuring the
PT   stability of HLA class I-peptide complexes on intact B cells
XX
PS   Example 3; Page 85; 109pp; English.
XX
CC   Peptides AAW39430-W39734 are used in a novel method for the selection of
CC   immunogenic T-cell peptide epitopes present in polypeptide antigens. The
CC   method involves the identification of peptide sequences capable of
CC   binding to an HLA (human leukocyte antigen) class I molecule and
CC   measuring the binding of this epitope peptide to the HLA class I
CC   peptide. The stability of binding of the peptide and MHC (major
CC   histocompatablilty complex) class I molecule is measured on intact human
CC   B cells carrying the MHC molecule at their cell surfaces. The method can
CC   be used to select peptide epitopes for generating vaccines against a
CC   disease associated with the polypeptide, e.g. cancers or AIDS. The
CC   peptide epitopes are especially T-cell peptide epitopes with strong
CC   anti-tumour and anti-viral immune responses. Peptide AAW39723 is derived
CC   from the human carcino-embryonic antigen (CEA) and has the ability to
CC   bind to the human MHC Class I allele HLA-A2.1.
XX
SQ   Sequence   9 AA;

  Query Match              75.0%;  Score 36;  DB 18;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSGACLNL 9
              ||||| |||
Db          1 YLSGANLNL 9


RESULT 3
AAW77134
ID   AAW77134 standard; peptide; 9 AA.
XX
AC   AAW77134;
XX
DT   16-NOV-1998  (first entry)
XX
DE   CEA synthetic peptide epitope 1.
XX
KW   Tyrosinase; tyrosinase cytotoxic lymphocyte response;
KW   cytotoxic T lymphocyte; cysteine-depleted; melanoma.
XX
OS   Synthetic.
XX
PN   WO9833810-A2.
XX
PD   06-AUG-1998.
XX
PF   29-JAN-1998;   98WO-US01592.
XX
PR   30-JAN-1997;   97US-0037781.
XX
PA   (UYVI-) UNIV VIRGINIA PATENT FOUND.
XX
PI   Engelhard VH,  Hunt DF,  Kittlesen D,  Slingluff CL;
XX
DR   WPI; 1998-437388/37.
XX
PT   Disease specific immunogen  - comprises disease specific cytotoxic T
PT   lymphocyte epitope used to elicit melanoma specific CTL response
XX
PS   Disclosure; Page 27; 93pp; English.
XX
CC   The peptide epitope AAW77119-W77138 were created for human
CC   tumour-specific cytotoxic T lymphocyte response.  These peptides are are
CC   cysteine- depleted mutants of a native disease-specific CTL epitope.  The
CC   cysteine- depleted CTL epitopes elicit a stronger or more specific CTL
CC   response than the native epitope.  The epitopes can be used in a
CC   disease-specific immunogen to protect a mammal against disease in
CC   particular melanomas. The peptides may also be used to screen a sample
CC   for the presence of an antigen with the same epitope, or with a different
CC   cross-reactive epitope.
XX
SQ   Sequence   9 AA;

  Query Match              75.0%;  Score 36;  DB 19;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSGACLNL 9
              ||||| |||
Db          1 YLSGANLNL 9
```

```
RESULT 4
AAW70045
ID   AAW70045 standard; peptide; 9 AA.
XX
AC   AAW70045;
XX
DT   22-OCT-1998  (first entry)
XX
DE   CEA derived HLA-A2.1 binding peptide 2 (residues 605-613).
XX
KW   Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW   human leukocyte antigen; HLA; tumour associated antigen; cancer;
KW   antigen presenting cell; APC; immunogenic peptide; immune disorder;
KW   viral infection; AIDS; hepatitis; bacterial infection; malaria; CEA;
KW   fungal infection; tuberculosis; melanoma; carcinoembryonic antigen.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9833888-A1.
XX
PD   06-AUG-1998.
XX
PF   30-JAN-1998;   98WO-US01959.
XX
PR   31-JAN-1997;   97US-0036696.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Celis E,  Sette A,  Sidney J,  Southwood S,  Tsai V;
XX
DR   WPI; 1998-437445/37.
XX
PT   Production of antigen-specific cytotoxic T cells - by incubating
PT   immunogenic peptide(s) from antigen that binds class I major
PT   histocompatibility complex molecules with pre-treated antigen
PT   presenting cells
XX
PS   Example 6; Page 75; 104pp; English.
XX
CC   Sequences shown in AAW70044 to AAW70052 represent peptides derived from
CC   carcinoembryonic antigen (CEA). The peptides can bind to a human
CC   leukocyte antigen (HLA), HLA-A2.1 and are used to exemplify the method
CC   of invention of producing antigen-specific cytotoxic T cells (CTLs) in
CC   vitro. The method comprises contacting immunogenic peptides from an
CC   antigen that binds class I major histocompatibility complex (MHC)
CC   molecules with antigen presenting cells (APCs) pretreated with
CC   pretreatment growth factors, and incubating the APCs with purified CD8
CC   cells in the presence of at least 2 incubation growth factors, thereby
CC   producing antigen-specific CTLs. A method for specifically killing
CC   target cells in a human patient is also provided which comprises
CC   obtaining a fluid sample containing CTLs from a patient, contacting the
CC   cytotoxic T cells with APCs pretreated with pre-treatment growth
CC   factors, where the APCs comprise class I MHC molecules. The pretreated
CC   APCs are incubated with the cytotoxic growth factors, thereby producing
CC   activated CTLs which are contacted with a carrier to form a composition.
CC   The composition can then be administered to the patient. The activated
CC   CTLs can be used for treating cancers, immune disorders, viral
CC   infections, AIDS, hepatitis, bacterial infection, fungal infection,
CC   malaria or tuberculosis.
XX
SQ   Sequence   9 AA;

  Query Match             75.0%;  Score 36;  DB 19;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSGACLNL 9
              ||||| |||
Db          1 YLSGANLNL 9
```

```
RESULT 5
AAY47655
ID   AAY47655 standard; Peptide; 9 AA.
XX
AC   AAY47655;
XX
DT   01-DEC-1999  (first entry)
XX
DE   Immunogenic peptide having a human leukocyte antigen binding motif #2266.
XX
KW   Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW   immune response; T cell activation; major histocompatibility complex;
KW   cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW   prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW   vaccine; immunisation.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9945954-A1.
XX
PD   16-SEP-1999.
XX
PF   13-MAR-1998;   98WO-US05039.
XX
PR   13-MAR-1998;   98WO-US05039.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Kubo RT,  Sidney J,  Celis E,  Grey HM,  Southwood S;
XX
DR   WPI; 1999-551214/46.
XX
PT   New immunogenic peptides with HLA binding motif, useful in treatment
PT   and diagnosis of cancers and viral diseases  -
XX
PS   Claim 1; Page 118; 150pp; English.
XX
CC   AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
CC   having a human major histocompatibility complex (MHC) Class I (also
CC   known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC   peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC   HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC   response against the antigen from which the peptide is derived.
CC   Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC   normally induced by an antigen in the form of a peptide fragment bound
CC   to a HLA molecule, rather than the intact foreign antigen itself, and
CC   are particularly important in tumour rejection and in fighting viral
CC   infections. The peptides are therefore useful therapeutically to treat
CC   or prevent viral infections and cancers in mammals (especially humans)
CC   e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC   They can be administered as vaccines to elicit an immune response in
CC   individuals susceptible or otherwise at risk of viral infection or
CC   cancer, or used to treat chronic or acute conditions. They are also
CC   useful diagnostically, and can be used to induce a cytotoxic T cell
CC   response, by contacting a cytotoxic T cell with the peptide e.g. to
CC   produce CTLs ex vivo for infusion back into a patient. The
CC   polynucleotides encoding the immunogenic peptides are also useful
CC   therapeutically and for immunisation as above.
XX
SQ   Sequence   9 AA;

  Query Match             75.0%;  Score 36;  DB 20;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSGACLNL 9
              ||||| |||
Db          1 YLSGANLNL 9


RESULT 6
AAY09525
```

```
ID   AAY09525 standard; peptide; 9 AA.
XX
AC   AAY09525;
XX
DT   20-JUL-1999  (first entry)
XX
DE   Carcinoembryonic antigen peptide agonist CAP-1.
XX
KW   Carcinoembryonic antigen; CEA; human; agonist; antagonist;
KW   immune response; carcinoma; gastrointestinal; breast; pancreatic;
KW   bladder; ovarian; lung; prostatic; T cell proliferation; cancer;
KW   adoptive transfer therapy; autoimmune reaction; immunotherapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9919478-A1.
XX
PD   22-APR-1999.
XX
PF   22-SEP-1998;   98WO-US19794.
XX
PR   10-OCT-1997;   97US-0061589.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Barzaga E,  Schlom J,  Zaremba S;
XX
DR   WPI; 1999-326544/27.
XX
PT   Peptide agonists and antagonists of carcinoembryonal antigen
XX
PS   Claim 1; Page 53; 72pp; English.
XX
CC   The present invention describes peptides (A) that comprise agonists (Ia)
CC   or antagonists (Ib) of human carcinoembryonal antigen (CEA). (Ia) are
CC   used in vaccines to kill or inhibit carcinoma cells that express CEA or
CC   its epitopes, particularly for treating gastrointestinal, breast,
CC   pancreatic, bladder, ovarian, lung or prostatic carcinoma. They can also
CC   be used to proliferate T cells, e.g. from vaccinated subjects, for use
CC   in adoptive transfer therapy. (Ib) are used to inhibit CEA-specific
CC   immune responses, e.g. in vaccinated subjects, to prevent an autoimmune
CC   reaction to cancer immunotherapy (i.e. to prevent attack on normal but
CC   CEA-expressing cells). (Ia) are more active than native sequence (I) and
CC   generate a highly specific and systemic anti-CEA response. Cytotoxic T
CC   cells generated recognize both (Ia) and native CEA epitopes. The present
CC   sequence represents a specifically claimed example of (Ia).
XX
SQ   Sequence   9 AA;

  Query Match             75.0%;  Score 36;  DB 20;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSGACLNL 9
                ||||| |||
Db          1 YLSGANLNL 9


RESULT 7
AAY09526
ID   AAY09526 standard; peptide; 9 AA.
XX
AC   AAY09526;
XX
DT   20-JUL-1999  (first entry)
XX
DE   Carcinoembryonic antigen peptide agonist SEQ ID NO:2.
XX
KW   Carcinoembryonic antigen; CEA; human; agonist; antagonist;
KW   immune response; carcinoma; gastrointestinal; breast; pancreatic;
KW   bladder; ovarian; lung; prostatic; T cell proliferation; cancer;
KW   adoptive transfer therapy; autoimmune reaction; immunotherapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9919478-A1.
XX
PD   22-APR-1999.
XX
PF   22-SEP-1998;   98WO-US19794.
XX
PR   10-OCT-1997;   97US-0061589.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Barzaga E,  Schlom J,  Zaremba S;
XX
DR   WPI; 1999-326544/27.
XX
PT   Peptide agonists and antagonists of carcinoembryonal antigen
XX
PS   Claim 5; Page 53; 72pp; English.
XX
CC   The present invention describes peptides (A) that comprise agonists (Ia)
CC   or antagonists (Ib) of human carcinoembryonal antigen (CEA). (Ia) are
CC   used in vaccines to kill or inhibit carcinoma cells that express CEA or
CC   its epitopes, particularly for treating gastrointestinal, breast,
CC   pancreatic, bladder, ovarian, lung or prostatic carcinoma. They can also
CC   be used to proliferate T cells, e.g. from vaccinated subjects, for use
CC   in adoptive transfer therapy. (Ib) are used to inhibit CEA-specific
CC   immune responses, e.g. in vaccinated subjects, to prevent an autoimmune
CC   reaction to cancer immunotherapy (i.e. to prevent attack on normal but
CC   CEA-expressing cells). (Ia) are more active than native sequence (I) and
CC   generate a highly specific and systemic anti-CEA response. Cytotoxic T
CC   cells generated recognize both (Ia) and native CEA epitopes. The present
CC   sequence represents a specifically claimed example of (Ia).
XX
SQ   Sequence   9 AA;

  Query Match             75.0%;  Score 36;  DB 20;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSGACLNL 9
                ||||| |||
Db          1 YLSGADLNL 9


RESULT 8
AAB13749
ID   AAB13749 standard; peptide; 9 AA.
XX
AC   AAB13749;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Peptide fragment # 1 from human CEA.
XX
KW   Human; T-cell; immune response; antigen; epitope; B7 family molecule;
KW   Leukocyte function-associated antigen-3; LFA-3;
KW   Intercellular adhesion molecule-1; ICAM-1; vaccine; immunotherapy;
KW   colon polyp; Crohn's disease; ulcerative colitis; breast lesion;
KW   tumour; CEA.
XX
OS   Homo sapiens.
XX
PN   WO200034494-A1.
XX
PD   15-JUN-2000.
XX
PF   12-NOV-1999;   99WO-US26866.
XX
PR   09-DEC-1998;   98US-0111582.
XX
```

```
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (THER-) THERION BIOLOGICS CORP.
XX
PI   Schlom J,  Hodge J,  Panicali D;
XX
DR   WPI; 2000-431307/37.
XX
PT   Novel recombinant vector useful as immunogens and vaccines for
PT   stimulating and enhancing immunological responses to target cells and
PT   antigens expresses multiple co-stimulatory molecules such as B7-1,
PT   LFA-3, ICAM-1  -
XX
PS   Claim 18; Page 35; 188pp; English.
XX
CC   Costimulatory molecules have important roles in T-cell activation and
CC   therefore the immune response. The present invention relates to
CC   recombinant vectors which comprise of foreign nucleic acid sequences
CC   encoding at least three costimulatory molecules: a B7 family molecule,
CC   Leukocyte function-associated antigen-3 (LFA-3, human CD58) and
CC   Intercellular adhesion molecule-1 (ICAM-1, CD54) and optionally a foreign
CC   gene encoding a target antigen or immunological epitope. The present
CC   sequence is one such target antigen used in the present invention. The
CC   present sequence is a tumour-associated antigen. The vector of the
CC   present invention would be useful for providing an enhanced immune
CC   response to the present target antigen. The vector of the present
CC   invention may therefore be useful in immunotherapy for treating or
CC   preventing diseases caused by viruses, bacteria, protozoans, parasites,
CC   premalignant cells and tumour cells. The recombinant vector can be used
CC   to treat or prevent preneoplastic or hyperplastic states such as colon
CC   polyps, Crohn's disease, ulcerative colitis and breast lesions.
XX
SQ   Sequence   9 AA;

  Query Match             75.0%;  Score 36;  DB 21;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSGACLNL 9
              ||||| |||
Db          1 YLSGANLNL 9


RESULT 9
AAB13750
ID   AAB13750 standard; peptide; 9 AA.
XX
AC   AAB13750;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Peptide fragment # 2 from human CEA.
XX
KW   Human; T-cell; immune response; antigen; epitope; B7 family molecule;
KW   Leukocyte function-associated antigen-3; LFA-3;
KW   Intercellular adhesion molecule-1; ICAM-1; vaccine; immunotherapy;
KW   colon polyp; Crohn's disease; ulcerative colitis; breast lesion;
KW   tumour; CEA.
XX
OS   Homo sapiens.
XX
PN   WO200034494-A1.
XX
PD   15-JUN-2000.
XX
PF   12-NOV-1999;   99WO-US26866.
XX
PR   09-DEC-1998;   98US-0111582.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (THER-) THERION BIOLOGICS CORP.
XX
PI   Schlom J,  Hodge J,  Panicali D;
XX
```

```
DR   WPI; 2000-431307/37.
XX
PT   Novel recombinant vector useful as immunogens and vaccines for
PT   stimulating and enhancing immunological responses to target cells and
PT   antigens expresses multiple co-stimulatory molecules such as B7-1,
PT   LFA-3, ICAM-1  -
XX
PS   Claim 18; Page 35; 188pp; English.
XX
CC   Costimulatory molecules have important roles in T-cell activation and
CC   therefore the immune response. The present invention relates to
CC   recombinant vectors which comprise of foreign nucleic acid sequences
CC   encoding at least three costimulatory molecules: a B7 family molecule,
CC   Leukocyte function-associated antigen-3 (LFA-3, human CD58) and
CC   Intercellular adhesion molecule-1 (ICAM-1, CD54) and optionally a foreign
CC   gene encoding a target antigen or immunological epitope. The present
CC   sequence is one such target antigen used in the present invention. The
CC   present sequence is a tumour-associated antigen. The vector of the
CC   present invention would be useful for providing an enhanced immune
CC   response to the present target antigen. The vector of the present
CC   invention may therefore be useful in immunotherapy for treating or
CC   preventing diseases caused by viruses, bacteria, protozoans, parasites,
CC   premalignant cells and tumour cells. The recombinant vector can be used
CC   to treat or prevent preneoplastic or hyperplastic states such as colon
CC   polyps, Crohn's disease, ulcerative colitis and breast lesions.
XX
SQ   Sequence   9 AA;

  Query Match             75.0%;  Score 36;  DB 21;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSGACLNL 9
              ||||| |||
Db          1 YLSGADLNL 9

RESULT 10
AAB82776
ID   AAB82776 standard; Protein; 9 AA.
XX
AC   AAB82776;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Carcinoembryonic antigen peptide.
XX
KW   Telomerase reverse transcriptase; hTRT; human;
KW   cytotoxic T lymphocyte; major histocompatibility complex; cancer;
KW   tumour; human leucocyte antigen; HLA-A2.1; vaccine;
KW   carcinoembryonic antigen.
XX
OS   Homo sapiens.
XX
PN   WO200160391-A1.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05143.
XX
PR   15-FEB-2000; 2000US-0182685.
PR   15-FEB-2001; 2001US-0182685.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Zanetti M;
XX
DR   WPI; 2001-536552/59.
XX
PT   Vaccine for initiating and enhancing a cytotoxic T lymphocyte response,
PT   for treating cancers or tumours or for inducing immune response against
PT   tumours, comprises a telomerase reverse transcriptase peptide    -
XX
```

```
PS   Example 1; Page 12; 52pp; English.
XX
CC   The present sequence is that of a carcinoembryonic antigen peptide
CC   comprising amino acid residues 571-579.  The peptide was used as a
CC   reference peptide in comparison with human telomerase reverse
CC   transcriptase (hTRT) HLA-A2.1+ restricted peptide p540 (see
CC   AAB82772) in a HLA-A2.1 binding/stabilisation assay.  The induction
CC   of CTL responses in vitro and in vivo, and the susceptibility to
CC   lysis of tumour cells of various origins by hTRT CTL suggest that
CC   hTRT could serve as a universal cancer vaccine for humans.  A
CC   claimed universal vaccine for treating tumours of any origin
CC   comprises at least 1 hTRT peptide.  The peptide is 7-15 amino
CC   acid residues in length and may be modified to enhance binding to
CC   the major histocompatibility complex.  Also claimed is a method for
CC   inducing and enhancing a CTL response against cancer cells, involving
CC   harvesting blood leucocytes, pulsing with hTRT, and contacting
CC   cancer cells with the pulsed leucocytes.  A method for targeting
CC   CTL to tumour cells is also claimed, and involves administering a
CC   hTRT peptide to a mammal, especially a cancer patient.
XX
SQ   Sequence   9 AA;

  Query Match             75.0%;  Score 36;  DB 22;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSGACLNL 9
                ||||| |||
Db          1 YLSGANLNL 9


RESULT 11
AAE05123
ID   AAE05123 standard; peptide; 9 AA.
XX
AC   AAE05123;
XX
DT   18-SEP-2001  (first entry)
XX
DE   Carcinoembryonic antigen (CEA) peptide, CAP-1.
XX
KW   Tumour-associated antigen; TAA; cytostatic; vaccine; gene therapy;
KW   immune response; tetanus toxoid; TT; diphtheria toxoid; DT; prophylactic;
KW   cancer; therapeutic; carcinoembryonic antigen; CEA.
XX
OS   Unidentified.
XX
PN   WO200149317-A2.
XX
PD   12-JUL-2001.
XX
PF   05-JAN-2001; 2001WO-CA00005.
XX
PR   05-JAN-2000; 2000US-0174587.
XX
PA   (AVET ) AVENTIS PASTEUR LTD.
XX
PI   Emtage P,  Barber BH,  Sambhara S,  Sia CDY;
XX
DR   WPI; 2001-441790/47.
XX
PT   Enhancing immune response to antigen such as tumor antigen for treating
PT   cancer in an animal involves administering an inducing agent to the
PT   animal followed by administering inducing agent-antigen mixture  -
XX
PS   Example 2; Page 31; 62pp; English.
XX
CC   The invention relates to a method of enhancing an immune response against
CC   tumour-associated antigens (TAAs), such as GP100 and carcinoembryonic
CC   antigen (CEA) in an animal. The method involves priming of the animal
CC   with an inducing agent such as tetanus toxoid (TT) or diphtheria toxoid
CC   (DT), subsequently followed by administration of an inducing agent-
CC   antigen mixture. The method provides the enhancement or augmentation of
CC   the immune response to the antigen and/or improves a vaccination protocol
CC   by allowing use of less antigen. The immunisation of the animal with
CC   tumour-associated antigen is useful for the prophylactic or therapeutic
CC   treatment of cancer. The present sequence is carcinoembryonic antigen
CC   (CEA) peptide fragment related to the invention.
XX
SQ   Sequence   9 AA;

  Query Match             75.0%;  Score 36;  DB 22;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSGACLNL 9
                ||||| |||
Db          1 YLSGANLNL 9


RESULT 12
AAE05124
ID   AAE05124 standard; peptide; 9 AA.
XX
AC   AAE05124;
XX
DT   18-SEP-2001  (first entry)
XX
DE   Modified carcinoembryonic antigen (CEA) peptide, CAP-6D.
XX
KW   Tumour-associated antigen; TAA; cytostatic; vaccine; gene therapy;
KW   immune response; tetanus toxoid; TT; diphtheria toxoid; DT; prophylactic;
KW   cancer; therapeutic; carcinoembryonic antigen; CEA.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 6
FT                   /note= "Wild type Asn substituted with Asp"
XX
PN   WO200149317-A2.
XX
PD   12-JUL-2001.
XX
PF   05-JAN-2001; 2001WO-CA00005.
XX
PR   05-JAN-2000; 2000US-0174587.
XX
PA   (AVET ) AVENTIS PASTEUR LTD.
XX
PI   Emtage P,  Barber BH,  Sambhara S,  Sia CDY;
XX
DR   WPI; 2001-441790/47.
XX
PT   Enhancing immune response to antigen such as tumor antigen for treating
PT   cancer in an animal involves administering an inducing agent to the
PT   animal followed by administering inducing agent-antigen mixture  -
XX
PS   Example 2; Page 31; 62pp; English.
XX
CC   The invention relates to a method of enhancing an immune response against
CC   tumour-associated antigens (TAAs), such as GP100 and carcinoembryonic
CC   antigen (CEA) in an animal. The method involves priming of the animal
CC   with an inducing agent such as tetanus toxoid (TT) or diphtheria toxoid
CC   (DT), subsequently followed by administration of an inducing agent-
CC   antigen mixture. The method provides the enhancement or augmentation of
CC   the immune response to the antigen and/or improves a vaccination protocol
CC   by allowing use of less antigen. The immunisation of the animal with
CC   tumour-associated antigen is useful for the prophylactic or therapeutic
CC   treatment of cancer. The present sequence is modified carcinoembryonic
CC   antigen (CEA) peptide fragment related to the invention.
XX
SQ   Sequence   9 AA;

  Query Match             75.0%;  Score 36;  DB 22;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
```

```
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSGACLNL 9
              ||||| |||
Db          1 YLSGADLNL 9


RESULT 13
AAB97818
ID   AAB97818 standard; Peptide; 9 AA.
XX
AC   AAB97818;
XX
DT   08-AUG-2001  (first entry)
XX
DE   Carcinoembryonic antigen (CEA) modified antigen SEQ ID NO:113.
XX
KW   Virus; adenovirus; poxvirus; alphavirus; immune response; gp100;
KW   tumour antigen; CEA; carcinoembryonic antigen; immunostimulant;
KW   cytostatic; immunotherapy; interferon-gamma; IFN-gamma; cancer.
XX
OS   Unidentified.
XX
PN   WO200130382-A1.
XX
PD   03-MAY-2001.
XX
PF   20-OCT-2000; 2000WO-CA01253.
XX
PR   22-OCT-1999;   99US-0160879.
PR   07-AUG-2000; 2000US-0223325.
XX
PA   (AVET ) AVENTIS PASTEUR LTD.
XX
PI   Berinstein N,  Tartaglia J,  Moingeon P,  Barber B;
XX
DR   WPI; 2001-308587/32.
XX
PT   Inducing immune response to tumor antigen, useful in immunotherapy of
PT   cancer, by administering the antigen to a lymphatic site  -
XX
PS   Claim 19; Page 9; 60pp; English.
XX
CC   The present invention describes a method for inducing an immune response,
CC   in an animal, to a tumour antigen (Ag) comprising administering Ag, or
CC   nucleic acid (I) that encodes it, to a lymphatic site. Cynomolgus monkeys
CC   (Macaca fascicularis) were injected with a modified form of gp100 antigen
CC   (a) into the left inguinal lymph node or (b) subcutaneously. Both animals
CC   of (a) developed a cell-mediated response (indicated by production of
CC   interferon-gamma from T lymphocytes when exposed to gp100 peptides), but
CC   only 2 of 4 animals of (b) did so. Also animals in (a) produced a far
CC   greater antibody response to gp100. The method is used in immunotherapy
CC   of a wide range of cancers through induction of a specific immune
CC   response (humoral and cellular) against the tumour antigens. When
CC   administered to a lymphatic site, Ag (or (I)) induces a stronger immune
CC   response than administration by other routes and may also break tolerance
CC   to Ag. AAB97708 and AAB97709 represent gp100 epitopes; AAB97710 to
CC   AAB97815 represent peptides derived from gp100 which stimulate interferon
CC   (IFN)-gamma production; AAH20120 encodes the modified gp100 protein given
CC   in AAB97816; AAH20121 encodes the modified carcinoembryonic antigen (CEA)
CC   protein given in AAB97817; and AAB97818 represents a CEA modified antigen
CC   peptide, all of which are used in the exemplification of the present
CC   invention.
XX
SQ   Sequence   9 AA;

  Query Match             75.0%;  Score 36;  DB 22;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSGACLNL 9
              ||||| |||
Db          1 YLSGADLNL 9
```

```
RESULT 14
AAE02673
ID   AAE02673 standard; peptide; 9 AA.
XX
AC   AAE02673;
XX
DT   06-AUG-2001  (first entry)
XX
DE   Human CEA epitopic peptide.
XX
KW   Human; cytostatic; antibacterial; antifungal; gene therapy; vaccine;
KW   antiviral; tumour; epitope; glycoprotein; hepatitis B virus; HBV;
KW   immune response; CTL; cytotoxic T lymphocyte; CEA; HLA;
KW   human leucocyte antigen.
XX
OS   Homo sapiens.
XX
PN   WO200127291-A1.
XX
PD   19-APR-2001.
XX
PF   29-SEP-2000; 2000WO-EP09902.
XX
PR   12-OCT-1999;   99US-0158356.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Firat H,  Lemonnier F,  Langlade-demoyen P;
XX
DR   WPI; 2001-282038/29.
XX
PT   New polynucleotide comprising at least one viral, fungal, bacterial, or
PT   tumor epitope of an antigen, capable of inducing a cellular response -
XX
PS   Example 1; Page 23; 70pp; English.
XX
CC   The invention relates to polynucleotide containing at least a part of
CC   the coding sequence of the middle glycoprotein of hepatitis B virus
CC   (HBV) in which is inserted a DNA sequence coding for an epitope
CC   comprising at least one viral, fungal, bacterial, or tumour epitope of
CC   an antigen, capable of inducing a cellular response. Nucleic acids and
CC   compositions of the invention are useful for inducing in vivo a CTL
CC   (cytotoxic T lymphocyte) response against several epitopes of one
CC   or more, bacterial, viral, fungal, or tumour antigens. A composition
CC   of the invention produces an immune response against HIV antigen and
CC   are used in the production of vaccines. The polynucleotides of the
CC   invention are also used in gene therapy. The present sequence is
CC   human CEA epitopic peptide. This peptide elicits strong
CC   HLA (human leucocyte antigen)-A2.1-restricted CTL response in mice.
XX
SQ   Sequence   9 AA;

  Query Match             75.0%;  Score 36;  DB 22;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSGACLNL 9
              ||||| |||
Db          1 YLSGANLNL 9


RESULT 15
AAE00463
ID   AAE00463 standard; peptide; 9 AA.
XX
AC   AAE00463;
XX
DT   19-JUN-2001  (first entry)
XX
DE   Human tumour CEA epitopic peptide.
XX
```

```
KW   Human; tumour epitope; cytostatic; immunostimulant; gene therapy;
KW   middle glycoprotein; Hepatitis B virus; HBV; cytotoxic response;
KW   immune response; cytotoxic T lymphocyte; CTL; CEA; HLA;
KW   human leucocyte antigen.
XX
OS   Homo sapiens.
XX
PN   WO200123577-A2.
XX
PD   05-APR-2001.
XX
PF   29-SEP-2000; 2000WO-EP09900.
XX
PR   30-SEP-1999;   99US-0156945.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Firat H,  Lemonnier F,  Langlade-demoyen P,  Michel M,  Suhrbier AA;
XX
DR   WPI; 2001-266164/27.
XX
PT   Novel polynucleotide having DNA sequence encoding tumor antigen epitope
PT   inserted in part of coding sequence of middle glycoprotein of hepatitis
PT   B virus, used to induce immune response against tumor-specific antigen
PT   -
XX
PS   Example 1; Page 13; 36pp; English.
XX
CC   The present invention relates to an isolated or purified polynucleotide
CC   containing a DNA sequence coding for at least one tumour epitope of a
CC   tumour antigen inserted into part of the coding sequence of the middle
CC   glycoprotein of the Hepatitis B virus (HBV). The polynucleotide is
CC   useful for optionally evaluating cytotoxic responses in the individual's
CC   lymphocyte population. It induces an immune response against at least
CC   one tumour specific antigen or tissue specific antigen. The vector
CC   comprising the polynucleotide induces in vivo, cellular and/or humoral
CC   immune response. The composition comprising the polynucleotide induces
CC   in vivo, cytotoxic T lymphocyte (CTL) against one or more antigens or
CC   epitopes present on the hybrid protein. The polynucleotide is also
CC   useful in gene therapy.
CC   The present sequence is a human tumour CEA epitopic peptide. This
CC   peptide elicits strong HLA (human leucocyte antigen)-A2.1-restricted
CC   CTL response in mice.
XX
SQ   Sequence   9 AA;

  Query Match             75.0%;  Score 36;  DB 22;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YLSGACLNL 9
              ||||| |||
Db          1 YLSGANLNL 9


Search completed: January 12, 2004, 14:28:24
Job time : 32.25 secs
```

```
                   GenCore version 5.1.6
         Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         January 12, 2004, 14:24:44 ; Search time 25.75 Seconds
                                             (without alignments)
                                             90.193 Million cell updates/sec

Title:          US-09-529-121A-3
Perfect score:  45
Sequence:       1 YLSGADINL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters:      775

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_23:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_virus:*
                13:  sp_vertebrate:*
                14:  sp_unclassified:*
                15:  sp_rvirus:*
                16:  sp_bacteriap:*
                17:  sp_archeap:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                    SUMMARIES
                %
Result        Query
  No.  Score  Match Length DB  ID                     Description
--------------------------------------------------------------------------
    1     18   40.0      8  3  Q9HDS4                 Q9hds4 aspergillus
    2     16   35.6      8  5  Q94695                 Q94695 physarum po
    3     16   35.6      9  1  Q50832                 Q50832 methanococc
    4     15   33.3      7 11   Q63480                 Q63480 rattus norv
    5     15   33.3      8  3  P87225                 P87225 saccharomyc
    6     15   33.3      8  4  Q9UMC7                 Q9umc7 homo sapien
    7     15   33.3      8  6  Q8WNS1                 Q8wns1 bos taurus
    8     14   31.1      8  2  Q9X3K1                 Q9x3k1 prochloroco
    9     14   31.1      8  4  Q8IUB8                 Q8iub8 homo sapien
   10     14   31.1      8 13   Q9PS69                 Q9ps69 gallus gall
   11     14   31.1      8 13   P82079                 P82079 limnodynast
   12     14   31.1      9  6  Q9TRS0                 Q9trs0 oryctolagus
   13     14   31.1      9 10   Q9FEC0                 Q9fec0 hordeum vul
   14     14   31.1      9 11   O35953                 O35953 mus musculu
   15     14   31.1      9 15   Q8UTD7                 Q8utd7 human immun
   16     13   28.9      7  2  Q8KMS3                 Q8kms3 klebsiella
   17     13   28.9      7 11  O55184                   O55184 rattus norv
   18     13   28.9      9  2  Q9R7E8                  Q9r7e8 escherichia
   19     13   28.9      9  2  Q44377                  Q44377 aeromonas t
   20     13   28.9      9  2  Q44468                  Q44468 aeromonas v
   21     13   28.9      9  2  Q43928                  Q43928 aeromonas p
   22     13   28.9      9  2  Q44001                  Q44001 aeromonas e
   23     13   28.9      9  2  P83222                  P83222 streptomyce
   24     13   28.9      9 13   Q8JJ12                  Q8jj12 carassius a
   25     13   28.9      9 13   Q8JJ18                  Q8jj18 danio aequi
   26     13   28.9      9 13   Q8JJ14                  Q8jj14 danio frank
   27     13   28.9      9 13   Q8JJ10                  Q8jj10 notropis ch
   28     13   28.9      9 13   Q8JJ16                  Q8jj16 danio albol
   29     12   26.7      7 15   Q07624                  Q07624 rous sarcom
   30     12   26.7      8  4  Q15901                  Q15901 homo sapien
   31     12   26.7      8  4  Q9Y4J3                  Q9y4j3 homo sapien
   32     12   26.7      8  7  Q95213                  Q95213 oryctolagus
   33     12   26.7      8 13   Q91098                  Q91098 manorina me
   34     12   26.7      8 13   Q90498                  Q90498 erythrura g
   35     12   26.7      9  2  Q9R635                  Q9r635 chlamydia t
   36     12   26.7      9  4  Q16220                  Q16220 homo sapien
   37     12   26.7      9  4  O95953                  O95953 homo sapien
   38     12   26.7      9  5  Q27396                  Q27396 babesia bov
   39     12   26.7      9  5  O96417                  O96417 drosophila
   40     12   26.7      9  6  Q28112                  Q28112 bos taurus
   41     12   26.7      9  8  Q94NB1                  Q94nb1 microcebus
   42     12   26.7      9  8  Q94NB2                  Q94nb2 microcebus
   43     12   26.7      9  8  Q94NA9                  Q94na9 daubentonia
   44     12   26.7      9  8  Q94NB0                  Q94nb0 microcebus
   45     12   26.7      9 10   Q8VZZ3                  Q8vzz3 pinus radia

                                 ALIGNMENTS


RESULT 1
Q9HDS4
ID   Q9HDS4      PRELIMINARY;      PRT;     8 AA.
AC   Q9HDS4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   TrpC polyprotein (Fragment).
GN   TRPC.
OS   Aspergillus flavus.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX   NCBI_TaxID=5059;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A55;
RA   Geiser D.M., Dorner J.W., Horn B.W., Taylor J.W.;
RT   "The phylogenetics of mycotoxin and sclerotium production in
RT   Aspergillus flavus and Aspergillus oryzae.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF261861; AAG16135.1; -.
KW   Polyprotein.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  807 MW;  F3B2C72AB5B87DD6 CRC64;

  Query Match             40.0%;  Score 18;  DB 3;  Length 8;
  Best Local Similarity   33.3%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSGADI 7
              ::|:|:
Db          1 MAGSDL 6


RESULT 2
Q94695
ID   Q94695      PRELIMINARY;      PRT;     8 AA.
AC   Q94695;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
```

```
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   Actin (Fragment).
GN   ARDC.
OS   Physarum polycephalum (Slime mold).
OC   Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physarida;
OC   Physarum.
OX   NCBI_TaxID=5791;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96182101; PubMed=8622700;
RA   Benard M., Lagnel C., Pallotta D., Pierron G.;
RT   "Mapping of a replication origin within the promoter region of two
RT   unlinked, abundantly transcribed actin genes of Physarum
RT   polycephalum.";
RL   Mol. Cell. Biol. 16:968-976(1996).
DR   EMBL; M73459; AAB03706.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  878 MW;  F4C6C2CAAB187B16 CRC64;

  Query Match             35.6%;  Score 16;  DB 5;  Length 8;
  Best Local Similarity   33.3%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 LSGADI 7
              : | |:
Db          1 MEGEDV 6


RESULT 3
Q50832
ID   Q50832      PRELIMINARY;      PRT;     9 AA.
AC   Q50832;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1996 (TrEMBLrel. 01, Last annotation update)
DE   Intergenic AT-rich DNA sequence (Fragment).
OS   Methanococcus voltae.
OC   Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC   Methanococcaceae; Methanococcus.
OX   NCBI_TaxID=2188;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85230552; PubMed=4006907;
RA   Bollschweiler C., Kuehn R., Klein A.;
RT   "Non-repetitive AT-rich sequences are found in intergenic regions of
RT   Methanococcus voltae DNA.";
RL   EMBO J. 4:805-809(1985).
DR   EMBL; X02518; CAA26355.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1087 MW;  99ED005DC404405A CRC64;

  Query Match             35.6%;  Score 16;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DIN 8
              |||
Db          2 DIN 4


RESULT 4
Q63480
ID   Q63480      PRELIMINARY;      PRT;     7 AA.
AC   Q63480;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TR4-NS orphan receptor (Fragment).
GN   TR4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
```

```
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96198747; PubMed=8612486;
RA   Yoshikawa T., Makino S., Gao X.M., Xing G.Q., Chuang D.M.,
RA   Detera-Wadleigh S.D.;
RT   "Splice variants of rat TR4 orphan receptor: differential expression
RT   of novel sequences in the 5'-untranslated region and C-terminal
RT   domain.";
RL   Endocrinology 137:1562-1571(1996).
DR   EMBL; U59125; AAB02827.1; -.
KW   Receptor.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  758 MW;  672AA87864005350 CRC64;

  Query Match             33.3%;  Score 15;  DB 11;  Length 7;
  Best Local Similarity   33.3%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 LSGADI 7
              : | |:
Db          2 IRGGDL 7


RESULT 5
P87225
ID   P87225      PRELIMINARY;      PRT;     8 AA.
AC   P87225;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   GIN11 protein (Fragment).
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wedler H., Wedler E., Scharfe M., Wambutt R.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   MIPS;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z73169; CAA97518.2; -.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  1019 MW;  4E21A9C449D5B73B CRC64;

  Query Match             33.3%;  Score 15;  DB 3;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLS 3
              |||
Db          1 YLS 3


RESULT 6
Q9UMC7
ID   Q9UMC7      PRELIMINARY;      PRT;     8 AA.
AC   Q9UMC7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   SHMT protein (Fragment).
GN   SHMT.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Chave K.J., Snell K., Sanders P.G.;
```

```
RT   "Isolation and characterisation of human genomic sequences encoding
RT   cytosolic serine hydroxymethyltransferase.";
RL   Biochem. Soc. Trans. 25:53-53(1997).
DR   EMBL; Y14492; CAB54844.1; -.
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  868 MW;  7C205721E44AB5B8 CRC64;

  Query Match             33.3%;  Score 15;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GADINL 9
              |:| :|
Db          1 GSDNHL 6


RESULT 7
Q8WNS1
ID   Q8WNS1      PRELIMINARY;      PRT;     8 AA.
AC   Q8WNS1;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   X-linked zinc finger protein (Fragment).
GN   ZFX.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Poloumienko A., Blecher S.;
RT   "Comparison between intron-exon structures in ZFX and ZFY genes.";
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF045782; AAL58190.1; -.
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  904 MW;  DF1DC2C4472AAB1A CRC64;

  Query Match             33.3%;  Score 15;  DB 6;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DINL 9
              |:|:
Db          3 DLNV 6


RESULT 8
Q9X3K1
ID   Q9X3K1      PRELIMINARY;      PRT;     8 AA.
AC   Q9X3K1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   Cytochrome b (Fragment).
GN   PETB.
OS   Prochlorococcus sp.
OC   Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC   Prochlorococcus.
OX   NCBI_TaxID=1220;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Urbach E., Chisholm S.W.;
RT   "Genetic diversity in Prochlorococcus populations flow cytometrically
RT   sorted from the Sargasso Sea and Gulf Stream.";
RL   Limnol. Oceanog. 43:1615-1630(1998).
DR   EMBL; AF070193; AAD23233.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  799 MW;  10376865B72866D3 CRC64;
```

```
  Query Match             31.1%;  Score 14;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |||
Db          4 LSG 6


RESULT 9
Q8IUB8
ID   Q8IUB8      PRELIMINARY;      PRT;     8 AA.
AC   Q8IUB8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CD95 antigen (Fragment).
GN   CD95.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22404279; PubMed=12516573;
RA   Kurth J., Perniok A., Schmitz R., Iking-Konert C., Chiorazzi N.,
RA   Thompson K.M., Winkler T., Rajewsky K., Kueppers R.;
RT   "Lack of deleterious somatic mutations in the CD95 gene of
RT   plasmablasts from systemic lupus erythematosus patients and
RT   autoantibody-producing cell lines.";
RL   Eur. J. Immunol. 32:3785-3792(2002).
DR   EMBL; AJ509178; CAD48928.1; -.
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  846 MW;  34B724405DC2D1AB CRC64;

  Query Match             31.1%;  Score 14;  DB 4;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 INL 9
              |||
Db          5 INL 7


RESULT 10
Q9PS69
ID   Q9PS69      PRELIMINARY;      PRT;     8 AA.
AC   Q9PS69;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   LOW density lipoprotein receptor-related protein (Fragment).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92011685; PubMed=1918027;
RA   Stifani S., Barber D.L., Aebersold R., Steyrer E., Shen X., Nimpf J.,
RA   Schneider W.J.;
RT   "The laying hen expresses two different low density lipoprotein
RT   receptor-related proteins.";
RL   J. Biol. Chem. 266:19079-19087(1991).
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  846 MW;  C007272DD865BAAA CRC64;

  Query Match             31.1%;  Score 14;  DB 13;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          3 SGA 5
              |||
Db          3 SGA 5

RESULT 11
P82079
ID   P82079      PRELIMINARY;      PRT;     8 AA.
AC   P82079;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DYNASTIN 1.
OS   Limnodynastes interioris (Giant banjo frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Limnodynastinae; Limnodynastes.
OX   NCBI_TaxID=30362;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=TIBIAL GLAND;
RA   Raftery M.J., Bradford A.M., Bowie J.H., Wallace J.C., Tyler M.J.;
RT   "Peptides from Australian frogs. The structure of the dynastins from
RT   the banjo frogs Limnodynastes interioras, Limnodynastes dumerilii and
RT   Limnodynastes terraereginae.";
RL   Aust. J. Chem. 46:833-842(1993).
CC   -!- MASS SPECTROMETRY: MW=729; METHOD=FAB.
KW   Amphibian skin.
SQ   SEQUENCE   8 AA;  729 MW;  7C28772865B72728 CRC64;

  Query Match             31.1%;  Score 14;  DB 13;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |||
Db          3 LSG 5

RESULT 12
Q9TRS0
ID   Q9TRS0      PRELIMINARY;      PRT;     9 AA.
AC   Q9TRS0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Calcyclin-associated protein, CAP50=CA2+/phospholipid-binding protein
DE   L-7 fragment (Fragment).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92250478; PubMed=1533622;
RA   Tokumitsu H., Mizutani A., Minami H., Kobayashi R., Hidaka H.;
RT   "A calcyclin-associated protein is a newly identified member of the
RT   Ca2+/phospholipid-binding proteins, annexin family.";
RL   J. Biol. Chem. 267:8919-8924(1992).
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1010 MW;  64E419C44865B72B CRC64;

  Query Match             31.1%;  Score 14;  DB 6;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |||
Db          3 LSG 5
```

```
RESULT 13
Q9FEC0
ID   Q9FEC0      PRELIMINARY;      PRT;     9 AA.
AC   Q9FEC0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Mla1uORF (MLA13uORF 2de).
GN   MLA1.
OS   Hordeum vulgare (Barley).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Hordeum.
OX   NCBI_TaxID=4513;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Algerian;
RA   Zhou F., Kurth J., Wei F., Elliott C., Vale G., Yahiaoui N.,
RA   Keller B., Somerville S., Wise R., Schulze-lefert P.;
RT   "Cell-autonomous Expression of Barley Mla1 Confers Race-specific
RT   Resistance to the Powdery Mildew Fungus via a Rar1 Independent
RT   Signaling Pathway.";
RL   Plant Cell 0:0-0(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Halterman D.A., Wei F., Wise R.P.;
RT   "Powdery mildew-induced Mla mRNAs are alternatively spliced and
RT   contain multiple upstream open reading frames.";
RL   Plant Physiol. 0:0-0(2003).
DR   EMBL; AY009939; AAG37357.1; -.
DR   EMBL; AY009938; AAG37355.1; -.
DR   EMBL; AF523682; AAO16013.1; -.
DR   EMBL; AF523683; AAO16016.1; -.
SQ   SEQUENCE   9 AA;  1163 MW;  473E2440573B5337 CRC64;

  Query Match             31.1%;  Score 14;  DB 10;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 INL 9
              |||
Db          6 INL 8

RESULT 14
O35953
ID   O35953      PRELIMINARY;      PRT;     9 AA.
AC   O35953;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Truncated voltage-gated sodium channel alpha subunit (Fragment).
GN   SCN8A.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RIII;
RX   MEDLINE=97442476; PubMed=9295353;
RA   Plummer N.W., McBurney M.W., Meisler M.H.;
RT   "Alternative splicing of the sodium channel SCN8A predicts a truncated
RT   two-domain protein in fetal brain and non-neuronal cells.";
RL   J. Biol. Chem. 272:24008-24015(1997).
DR   EMBL; U97672; AAB80914.1; -.
DR   MGD; MGI:103169; Scn8a.
KW   Ionic channel.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  898 MW;  22D92865B735B737 CRC64;

  Query Match             31.1%;  Score 14;  DB 11;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
```

```
 Matches     3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |||
Db          5 LSG 7


RESULT 15
Q8UTD7
ID   Q8UTD7      PRELIMINARY;      PRT;     9 AA.
AC   Q8UTD7;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Vpu protein.
GN   VPU.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=00BW1471.27;
RA   Novitsky V.A., Smith U.R., Gilbert P., McLane M.F., Chigwedere P.,
RA   Williamson C., Ndung'u T., Klein I., Chang S.-Y., Peter T., Thior I.,
RA   Foley B.T., Gaolekwe S., Rybak N., Gaseitsiwe S., Vannberg F.,
RA   Marlink R., Lee T.-H., Essex M.;
RT   "HIV-1 subtype C molecular phylogeny: consensus sequence for an AIDS
RT   vaccine design.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF443091; AAL34712.1; -.
SQ   SEQUENCE   9 AA;  1102 MW;  188BD40B17272440 CRC64;

  Query Match             31.1%;  Score 14;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 INL 9
              |||
Db          2 INL 4


Search completed: January 12, 2004, 14:30:59
Job time : 25.75 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         January 12, 2004, 14:21:08 ; Search time 32.25 Seconds
                                             (without alignments)
                                             44.296 Million cell updates/sec

Title:          US-09-529-121A-3
Perfect score:  45
Sequence:       1 YLSGADINL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters:      179625

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_19Jun03:*
                1:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
                2:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
                3:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
                4:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
                5:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
                6:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
                7:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
                8:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
                9:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
                10:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
                11:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
                12:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
                13:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                14:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
                15:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
                16:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                17:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
                18:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                19:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
                20:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                21:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
                22:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
                23:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
                24:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                %
Result          Query
  No.   Score   Match Length DB  ID                      Description
-------------------------------------------------------------------------
    1      45   100.0      9  20  AAY09527                Carcinoembryonic a
    2      43    95.6      9  20  AAY09526                Carcinoembryonic a
    3      43    95.6      9  21  AAB13750                Peptide fragment #
    4      43    95.6      9  22  AAE05124                Modified carcinoem
    5      43    95.6      9  22  AAB97818                Carcinoembryonic a
    6      43    95.6      9  23  AAB47917                Modified CEA epito
    7      43    95.6      9  23  AAE19089                HLA-A24 restricted
    8      40    88.9      9  20  AAY09528                Carcinoembryonic a
    9      38    84.4      9  18  AAW39723                Human carcina-embr
   10      38    84.4      9  19  AAW77134                CEA synthetic pept
   11      38    84.4      9  19  AAW70045                CEA derived HLA-A2
   12      38    84.4      9  20  AAY47655                Immunogenic peptid
   13      38    84.4      9  20  AAY09525                Carcinoembryonic a
   14      38    84.4      9  21  AAB13749                Peptide fragment #
   15      38    84.4      9  22  AAB82776                Carcinoembryonic a
   16      38    84.4      9  22  AAE05123                Carcinoembryonic a
   17      38    84.4      9  22  AAE02673                Human CEA epitopic
   18      38    84.4      9  22  AAE00463                Human tumour CEA e
   19      38    84.4      9  23  AAE26805                Human HLA-A2.1 res
   20      38    84.4      9  23  ABG79073                Human CEA class I
   21      38    84.4      9  23  AAU95893                Immunogenic peptid
   22      38    84.4      9  23  AAE19088                HLA-A24 restricted
   23      35    77.8      9  21  AAY54173                HLA binding peptid
   24      35    77.8      9  22  AAU26560                Human Leukocyte An
   25      35    77.8      9  22  AAB99681                HLA A2 binding CTL
   26      34    75.6      9  20  AAY09529                Carcinoembryonic a
   27      32    71.1      9  22  AAB75854                Tumour associated
   28      31    68.9      9  17  AAW00680                Peptide comprising
   29      26    57.8      9  23  AAU82064                CEA antigenic pept
   30      25    55.6      7  20  AAY41846                Rheumatoid arthrit
   31      24    53.3      8  22  ABP18160                HIV B58 super moti
   32      24    53.3      9  19  AAW70078                B. stearothermophi
   33      24    53.3      9  22  ABP18161                HIV B58 super moti
   34      23    51.1      7  20  AAY41847                Rheumatoid arthrit
   35      23    51.1      7  22  AAB75084                Nucleotide-5'-phos
   36      23    51.1      9  11  AAR07966                Tryptic fragment T
   37      23    51.1      9  18  AAW38383                Synthetic pMEL17 p
   38      23    51.1      9  20  AAY47062                Immunogenic peptid
   39      23    51.1      9  23  AAE31162                Human sp100 peptid
   40      23    51.1      9  23  AAE31390                Human PM17 peptide
   41      23    51.1      9  24  ABJ38049                Human cytomegalovi
   42      23    51.1      9  24  ABR16172                Human cancer-relat
   43      23    51.1      9  24  ABR16174                Human cancer-relat
   44      23    51.1      9  24  ABR16188                Human cancer-relat
   45      23    51.1      9  24  ABR16428                Human cancer-relat


                              ALIGNMENTS


RESULT 1
AAY09527
ID   AAY09527 standard; peptide; 9 AA.
XX
AC   AAY09527;
XX
DT   20-JUL-1999  (first entry)
XX
DE   Carcinoembryonic antigen peptide agonist SEQ ID NO:3.
XX
KW   Carcinoembryonic antigen; CEA; human; agonist; antagonist;
KW   immune response; carcinoma; gastrointestinal; breast; pancreatic;
KW   bladder; ovarian; lung; prostatic; T cell proliferation; cancer;
KW   adoptive transfer therapy; autoimmune reaction; immunotherapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9919478-A1.
XX
PD   22-APR-1999.
XX
PF   22-SEP-1998;   98WO-US19794.
XX
PR   10-OCT-1997;   97US-0061589.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Barzaga E,  Schlom J,  Zaremba S;
XX
DR   WPI; 1999-326544/27.
XX
```

```
PT   Peptide agonists and antagonists of carcinoembryonal antigen
XX
PS   Claim 5; Page 53; 72pp; English.
XX
CC   The present invention describes peptides (A) that comprise agonists (Ia)
CC   or antagonists (Ib) of human carcinoembryonal antigen (CEA). (Ia) are
CC   used in vaccines to kill or inhibit carcinoma cells that express CEA or
CC   its epitopes, particularly for treating gastrointestinal, breast,
CC   pancreatic, bladder, ovarian, lung or prostatic carcinoma. They can also
CC   be used to proliferate T cells, e.g. from vaccinated subjects, for use
CC   in adoptive transfer therapy. (Ib) are used to inhibit CEA-specific
CC   immune responses, e.g. in vaccinated subjects, to prevent an autoimmune
CC   reaction to cancer immunotherapy (i.e. to prevent attack on normal but
CC   CEA-expressing cells). (Ia) are more active than native sequence (I) and
CC   generate a highly specific and systemic anti-CEA response. Cytotoxic T
CC   cells generated recognize both (Ia) and native CEA epitopes. The present
CC   sequence represents a specifically claimed example of (Ia).
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 45;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADINL 9
              |||||||||
Db          1 YLSGADINL 9


RESULT 2
AAY09526
ID   AAY09526 standard; peptide; 9 AA.
XX
AC   AAY09526;
XX
DT   20-JUL-1999  (first entry)
XX
DE   Carcinoembryonic antigen peptide agonist SEQ ID NO:2.
XX
KW   Carcinoembryonic antigen; CEA; human; agonist; antagonist;
KW   immune response; carcinoma; gastrointestinal; breast; pancreatic;
KW   bladder; ovarian; lung; prostatic; T cell proliferation; cancer;
KW   adoptive transfer therapy; autoimmune reaction; immunotherapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9919478-A1.
XX
PD   22-APR-1999.
XX
PF   22-SEP-1998;   98WO-US19794.
XX
PR   10-OCT-1997;   97US-0061589.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Barzaga E,  Schlom J,  Zaremba S;
XX
DR   WPI; 1999-326544/27.
XX
PT   Peptide agonists and antagonists of carcinoembryonal antigen
XX
PS   Claim 5; Page 53; 72pp; English.
XX
CC   The present invention describes peptides (A) that comprise agonists (Ia)
CC   or antagonists (Ib) of human carcinoembryonal antigen (CEA). (Ia) are
CC   used in vaccines to kill or inhibit carcinoma cells that express CEA or
CC   its epitopes, particularly for treating gastrointestinal, breast,
CC   pancreatic, bladder, ovarian, lung or prostatic carcinoma. They can also
CC   be used to proliferate T cells, e.g. from vaccinated subjects, for use
CC   in adoptive transfer therapy. (Ib) are used to inhibit CEA-specific
CC   immune responses, e.g. in vaccinated subjects, to prevent an autoimmune
CC   reaction to cancer immunotherapy (i.e. to prevent attack on normal but
CC   CEA-expressing cells). (Ia) are more active than native sequence (I) and
CC   generate a highly specific and systemic anti-CEA response. Cytotoxic T
CC   cells generated recognize both (Ia) and native CEA epitopes. The present
CC   sequence represents a specifically claimed example of (Ia).
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 20;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADINL 9
              ||||||:||
Db          1 YLSGADLNL 9


RESULT 3
AAB13750
ID   AAB13750 standard; peptide; 9 AA.
XX
AC   AAB13750;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Peptide fragment # 2 from human CEA.
XX
KW   Human; T-cell; immune response; antigen; epitope; B7 family molecule;
KW   Leukocyte function-associated antigen-3; LFA-3;
KW   Intercellular adhesion molecule-1; ICAM-1; vaccine; immunotherapy;
KW   colon polyp; Crohn's disease; ulcerative colitis; breast lesion;
KW   tumour; CEA.
XX
OS   Homo sapiens.
XX
PN   WO200034494-A1.
XX
PD   15-JUN-2000.
XX
PF   12-NOV-1999;   99WO-US26866.
XX
PR   09-DEC-1998;   98US-0111582.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (THER-) THERION BIOLOGICS CORP.
XX
PI   Schlom J,  Hodge J,  Panicali D;
XX
DR   WPI; 2000-431307/37.
XX
PT   Novel recombinant vector useful as immunogens and vaccines for
PT   stimulating and enhancing immunological responses to target cells and
PT   antigens expresses multiple co-stimulatory molecules such as B7-1,
PT   LFA-3, ICAM-1  -
XX
PS   Claim 18; Page 35; 188pp; English.
XX
CC   Costimulatory molecules have important roles in T-cell activation and
CC   therefore the immune response. The present invention relates to
CC   recombinant vectors which comprise of foreign nucleic acid sequences
CC   encoding at least three costimulatory molecules: a B7 family molecule,
CC   Leukocyte function-associated antigen-3 (LFA-3, human CD58) and
CC   Intercellular adhesion molecule-1 (ICAM-1, CD54) and optionally a foreign
CC   gene encoding a target antigen or immunological epitope. The present
CC   sequence is one such target antigen used in the present invention. The
CC   present sequence is a tumour-associated antigen. The vector of the
CC   present invention would be useful for providing an enhanced immune
CC   response to the present target antigen. The vector of the present
CC   invention may therefore be useful in immunotherapy for treating or
CC   preventing diseases caused by viruses, bacteria, protozoans, parasites,
CC   premalignant cells and tumour cells. The recombinant vector can be used
CC   to treat or prevent preneoplastic or hyperplastic states such as colon
CC   polyps, Crohn's disease, ulcerative colitis and breast lesions.
```

```
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 21;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADINL 9
              ||||||:||
Db          1 YLSGADLNL 9


RESULT 4
AAE05124
ID   AAE05124 standard; peptide; 9 AA.
XX
AC   AAE05124;
XX
DT   18-SEP-2001  (first entry)
XX
DE   Modified carcinoembryonic antigen (CEA) peptide, CAP-6D.
XX
KW   Tumour-associated antigen; TAA; cytostatic; vaccine; gene therapy;
KW   immune response; tetanus toxoid; TT; diphtheria toxoid; DT; prophylactic;
KW   cancer; therapeutic; carcinoembryonic antigen; CEA.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 6
FT                   /note= "Wild type Asn substituted with Asp"
XX
PN   WO200149317-A2.
XX
PD   12-JUL-2001.
XX
PF   05-JAN-2001; 2001WO-CA00005.
XX
PR   05-JAN-2000; 2000US-0174587.
XX
PA   (AVET ) AVENTIS PASTEUR LTD.
XX
PI   Emtage P,  Barber BH,  Sambhara S,  Sia CDY;
XX
DR   WPI; 2001-441790/47.
XX
PT   Enhancing immune response to antigen such as tumor antigen for treating
PT   cancer in an animal involves administering an inducing agent to the
PT   animal followed by administering inducing agent-antigen mixture  -
XX
PS   Example 2; Page 31; 62pp; English.
XX
CC   The invention relates to a method of enhancing an immune response against
CC   tumour-associated antigens (TAAs), such as GP100 and carcinoembryonic
CC   antigen (CEA) in an animal. The method involves priming of the animal
CC   with an inducing agent such as tetanus toxoid (TT) or diphtheria toxoid
CC   (DT), subsequently followed by administration of an inducing agent-
CC   antigen mixture. The method provides the enhancement or augmentation of
CC   the immune response to the antigen and/or improves a vaccination protocol
CC   by allowing use of less antigen. The immunisation of the animal with
CC   tumour-associated antigen is useful for the prophylactic or therapeutic
CC   treatment of cancer. The present sequence is modified carcinoembryonic
CC   antigen (CEA) peptide fragment related to the invention.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 22;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADINL 9
              ||||||:||
Db          1 YLSGADLNL 9
```

```
RESULT 5
AAB97818
ID   AAB97818 standard; Peptide; 9 AA.
XX
AC   AAB97818;
XX
DT   08-AUG-2001  (first entry)
XX
DE   Carcinoembryonic antigen (CEA) modified antigen SEQ ID NO:113.
XX
KW   Virus; adenovirus; poxvirus; alphavirus; immune response; gp100;
KW   tumour antigen; CEA; carcinoembryonic antigen; immunostimulant;
KW   cytostatic; immunotherapy; interferon-gamma; IFN-gamma; cancer.
XX
OS   Unidentified.
XX
PN   WO200130382-A1.
XX
PD   03-MAY-2001.
XX
PF   20-OCT-2000; 2000WO-CA01253.
XX
PR   22-OCT-1999;   99US-0160879.
PR   07-AUG-2000; 2000US-0223325.
XX
PA   (AVET ) AVENTIS PASTEUR LTD.
XX
PI   Berinstein N,  Tartaglia J,  Moingeon P,  Barber B;
XX
DR   WPI; 2001-308587/32.
XX
PT   Inducing immune response to tumor antigen, useful in immunotherapy of
PT   cancer, by administering the antigen to a lymphatic site  -
XX
PS   Claim 19; Page 9; 60pp; English.
XX
CC   The present invention describes a method for inducing an immune response,
CC   in an animal, to a tumour antigen (Ag) comprising administering Ag, or
CC   nucleic acid (I) that encodes it, to a lymphatic site. Cynomolgus monkeys
CC   (Macaca fascicularis) were injected with a modified form of gp100 antigen
CC   (a) into the left inguinal lymph node or (b) subcutaneously. Both animals
CC   of (a) developed a cell-mediated response (indicated by production of
CC   interferon-gamma from T lymphocytes when exposed to gp100 peptides), but
CC   only 2 of 4 animals of (b) did so. Also animals in (a) produced a far
CC   greater antibody response to gp100. The method is used in immunotherapy
CC   of a wide range of cancers through induction of a specific immune
CC   response (humoral and cellular) against the tumour antigens. When
CC   administered to a lymphatic site, Ag (or (I)) induces a stronger immune
CC   response than administration by other routes and may also break tolerance
CC   to Ag. AAB97708 and AAB97709 represent gp100 epitopes; AAB97710 to
CC   AAB97815 represent peptides derived from gp100 which stimulate interferon
CC   (IFN)-gamma production; AAH20120 encodes the modified gp100 protein given
CC   in AAB97816; AAH20121 encodes the modified carcinoembryonic antigen (CEA)
CC   protein given in AAB97817; and AAB97818 represents a CEA modified antigen
CC   peptide, all of which are used in the exemplification of the present
CC   invention.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 22;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADINL 9
              ||||||:||
Db          1 YLSGADLNL 9


RESULT 6
AAB47917
ID   AAB47917 standard; peptide; 9 AA.
```

```
XX
AC   AAB47917;
XX
DT   16-MAY-2002  (first entry)
XX
DE   Modified CEA epitope, CEA(6D).
XX
KW   CAP-1; epitope; carcinoembryonic antigen; CEA; agonist; immune response;
KW   carcinoma; gastrointestinal; breast; pancreatic; bladder; ovarian;
KW   lung; prostate; cancer.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 6
FT                   /label= N6D
XX
PN   WO200210379-A2.
XX
PD   07-FEB-2002.
XX
PF   27-JUL-2001; 2001WO-CA01092.
XX
PR   31-JUL-2000; 2000US-222043P.
XX
PA   (AVET ) AVENTIS PASTEUR LTD.
PA   (THER-) THERION BIOLOGICS.
PA   (USSH ) US NAT CANCER INST.
XX
PI   Berinstein N,  Tartaglia J,  Tine JA,  Panicali DL,  Gritz L;
PI   Schlom J;
XX
DR   WPI; 2002-206189/26.
XX
PT   Carcinoembryonic antigen agonist polypeptide for inducing an immune
PT   response in animal against antigen and for inhibiting an epitope
PT   antigen expressing carcinoma cell, comprises a modified antigen epitope
PT   -
XX
PS   Claim 1; Page 38; 69pp; English.
XX
CC   This sequence represents a modified CAP-1 epitope of carcinoembryonic
CC   antigen (CEA) which was used as part of the CEA agonist polypeptide of
CC   the invention.  The modification of position 6 of this peptide from Asp
CC   to Asn increases its immunogenicity.  The CEA agonist polypeptide of
CC   the invention, or DNA encoding it, are useful for:
CC   (i) inducing an immune response in an animal directed against a CEA
CC   protein or fragment, CEA agonist, a CEA epitope, a modified CEA epitope,
CC   cells expressing or binding a CEA protein or fragment; and
CC   (ii) inhibiting a CEA epitope expressing carcinoma cell, which is a
CC   gastrointestinal, breast, pancreatic, bladder, ovarian, lung or
CC   prostate carcinoma cell in a patient, hence is useful for manufacture
CC   of a medicament for the treatment of cancer.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 23;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADINL 9
              ||||||:||
Db          1 YLSGADLNL 9


RESULT 7
AAE19089
ID   AAE19089 standard; peptide; 9 AA.
XX
AC   AAE19089;
XX
DT   21-MAY-2002  (first entry)
XX
DE   HLA-A24 restricted target antigen CEA immunological epitope #3.
XX
KW   Human leukocyte antigen; HLA; pharmaceutical composition; target antigen;
KW   immunological epitope; replication-defective virus; RDV; immune response;
KW   chemotherapy; granulocyte-monocyte-colony stimulating factor; cytostatic;
KW   GM-CSF; MHC; major histocompatibility complex; tumour; head; pancreatic;
KW   neck; breast; prostate; colorectal; melanoma; myeloidysplastic syndrome;
KW   metastatic breast skin lesion; corticosteroid therapy; erythropoietin;
KW   cytopenia; neutropenia; vaccine; immunostimulant.
XX
OS   Homo sapiens.
XX
PN   WO200195919-A2.
XX
PD   20-DEC-2001.
XX
PF   15-JUN-2001; 2001WO-US19201.
XX
PR   15-JUN-2000; 2000US-211717P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (THER-) THERION BIOLOGICS CORP.
XX
PI   Schlom J,  Greiner JW,  Kass E,  Panicali D;
XX
DR   WPI; 2002-205852/26.
XX
PT   Composition for enhancing immune responses, particularly anti-tumor
PT   responses and treating neutropenia, cytopenia, comprises
PT   replication-defective virus encoding granulocyte-monocyte-colony
PT   stimulating factor  -
XX
PS   Claim 9; Page 15; 118pp; English.
XX
CC   The present invention relates to a pharmaceutical composition comprising
CC   a replication-defective virus (RDV) encoding granulocyte-monocyte-colony
CC   stimulating factor (GM-CSF). The invention is useful for enhancing cell-
CC   mediated or humoral immune response in an individual, by enhancing
CC   migration of APC expressing CD11c^+/I-Ab^+, major histocompatibility
CC   complex (MHC) class II, at an injection site, regional lymph node at a
CC   tumour site, APC proliferation or function, CD4^+T or CD8^+T cell
CC   activation, interleukin (IL)-2, interferon (IFN)-gamma or tumour necrosis
CC   factor (TNF)-alpha production or their combinations. The composition
CC   enhances an antigen-specific T-cell response in an individual to a target
CC   antigen or its immunological epitope and an anti-tumour response in an
CC   individual with a head tumour, neck, breast, pancreatic, prostate,
CC   colorectal or metastatic tumour or melanoma, or metastatic breast skin
CC   lesion. The invention is further useful for treating neutropenia
CC   resulting from chemotherapy, corticosteroid therapy, irradiation or an
CC   infection, by raising the neutrophil count to normal levels and for
CC   treating cytopenias in patients with myeloidysplastic syndrome in
CC   combination with erythropoietin, by increasing neutrophil count and
CC   erythroid precursors. The composition enhances immune response to
CC   vaccines such as DPT, Td, DtaP, Hib, DtaP-Hib, MMR, Hepatitis A,
CC   hepatitis B, lyme's disease, influenza, tetravalent meningococcal
CC   polysaccharide, pneumococcal polysaccharide, anthrax, cholera, plague,
CC   yellow fever and Bacillus Calmette-Guerin vaccine. The present sequence
CC   is human leukocyte antigen (HLA)-restricted target tumour antigen
CC   immunological epitope.
XX
SQ   Sequence   9 AA;

  Query Match             95.6%;  Score 43;  DB 23;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADINL 9
              ||||||:||
Db          1 YLSGADLNL 9


RESULT 8
AAY09528
```

```
ID   AAY09528 standard; peptide; 9 AA.
XX
AC   AAY09528;
XX
DT   20-JUL-1999  (first entry)
XX
DE   Carcinoembryonic antigen peptide agonist SEQ ID NO:4.
XX
KW   Carcinoembryonic antigen; CEA; human; agonist; antagonist;
KW   immune response; carcinoma; gastrointestinal; breast; pancreatic;
KW   bladder; ovarian; lung; prostatic; T cell proliferation; cancer;
KW   adoptive transfer therapy; autoimmune reaction; immunotherapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9919478-A1.
XX
PD   22-APR-1999.
XX
PF   22-SEP-1998;   98WO-US19794.
XX
PR   10-OCT-1997;   97US-0061589.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Barzaga E,  Schlom J,  Zaremba S;
XX
DR   WPI; 1999-326544/27.
XX
PT   Peptide agonists and antagonists of carcinoembryonal antigen
XX
PS   Claim 5; Page 53; 72pp; English.
XX
CC   The present invention describes peptides (A) that comprise agonists (Ia)
CC   or antagonists (Ib) of human carcinoembryonal antigen (CEA). (Ia) are
CC   used in vaccines to kill or inhibit carcinoma cells that express CEA or
CC   its epitopes, particularly for treating gastrointestinal, breast,
CC   pancreatic, bladder, ovarian, lung or prostatic carcinoma. They can also
CC   be used to proliferate T cells, e.g. from vaccinated subjects, for use
CC   in adoptive transfer therapy. (Ib) are used to inhibit CEA-specific
CC   immune responses, e.g. in vaccinated subjects, to prevent an autoimmune
CC   reaction to cancer immunotherapy (i.e. to prevent attack on normal but
CC   CEA-expressing cells). (Ia) are more active than native sequence (I) and
CC   generate a highly specific and systemic anti-CEA response. Cytotoxic T
CC   cells generated recognize both (Ia) and native CEA epitopes. The present
CC   sequence represents a specifically claimed example of (Ia).
XX
SQ   Sequence   9 AA;

  Query Match             88.9%;  Score 40;  DB 20;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADINL 9
              |||||:|||
Db          1 YLSGANINL 9

RESULT 9
AAW39723
ID   AAW39723 standard; peptide; 9 AA.
XX
AC   AAW39723;
XX
DT   11-JUN-1998  (first entry)
XX
DE   Human carcina-embryonic antigen (CEA) peptide (pos. 571-579).
XX
KW   T cell epitope; immune response; human leukocyte antigen; HLA Class I;
KW   vaccine; immunogenic; major histocompatablilty complex; MHC; B cell;
KW   disease; anti-tumour; anti-viral.
XX
```

```
OS   Homo sapiens.
XX
PN   WO9741440-A1.
XX
PD   06-NOV-1997.
XX
PF   28-APR-1997;   97WO-NL00229.
XX
PR   23-DEC-1996;   96EP-0203670.
PR   26-APR-1996;   96EP-0201145.
XX
PA   (UYLE-) RIJKSUNIV LEIDEN.
PA   (SCIS-) SCI SEED CAPITAL INVESTMENTS BV.
XX
PI   Kast WM,  Melief CJM,  Offringa R,  Toes REM,  Van Der Burg SH;
XX
DR   WPI; 1997-549891/50.
XX
PT   Method of selecting T cell peptide epitope(s) - by measuring the
PT   stability of HLA class I-peptide complexes on intact B cells
XX
PS   Example 3; Page 85; 109pp; English.
XX
CC   Peptides AAW39430-W39734 are used in a novel method for the selection of
CC   immunogenic T-cell peptide epitopes present in polypeptide antigens. The
CC   method involves the identification of peptide sequences capable of
CC   binding to an HLA (human leukocyte antigen) class I molecule and
CC   measuring the binding of this epitope peptide to the HLA class I
CC   peptide. The stability of binding of the peptide and MHC (major
CC   histocompatablilty complex) class I molecule is measured on intact human
CC   B cells carrying the MHC molecule at their cell surfaces. The method can
CC   be used to select peptide epitopes for generating vaccines against a
CC   disease associated with the polypeptide, e.g. cancers or AIDS. The
CC   peptide epitopes are especially T-cell peptide epitopes with strong
CC   anti-tumour and anti-viral immune responses. Peptide AAW39723 is derived
CC   from the human carcino-embryonic antigen (CEA) and has the ability to
CC   bind to the human MHC Class I allele HLA-A2.1.
XX
SQ   Sequence   9 AA;

  Query Match             84.4%;  Score 38;  DB 18;  Length 9;
  Best Local Similarity   77.8%;  Pred. No. 9.3e+05;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADINL 9
              |||||::||
Db          1 YLSGANLNL 9

RESULT 10
AAW77134
ID   AAW77134 standard; peptide; 9 AA.
XX
AC   AAW77134;
XX
DT   16-NOV-1998  (first entry)
XX
DE   CEA synthetic peptide epitope 1.
XX
KW   Tyrosinase; tyrosinase cytotoxic lymphocyte response;
KW   cytotoxic T lymphocyte; cysteine-depleted; melanoma.
XX
OS   Synthetic.
XX
PN   WO9833810-A2.
XX
PD   06-AUG-1998.
XX
PF   29-JAN-1998;   98WO-US01592.
XX
PR   30-JAN-1997;   97US-0037781.
XX
PA   (UYVI-) UNIV VIRGINIA PATENT FOUND.
```

```
XX
PI   Engelhard VH,  Hunt DF,  Kittlesen D,  Slingluff CL;
XX
DR   WPI; 1998-437388/37.
XX
PT   Disease specific immunogen  - comprises disease specific cytotoxic T
PT   lymphocyte epitope used to elicit melanoma specific CTL response
XX
PS   Disclosure; Page 27; 93pp; English.
XX
CC   The peptide epitope AAW77119-W77138 were created for human
CC   tumour-specific cytotoxic T lymphocyte response.  These peptides are are
CC   cysteine- depleted mutants of a native disease-specific CTL epitope.  The
CC   cysteine- depleted CTL epitopes elicit a stronger or more specific CTL
CC   response than the native epitope.  The epitopes can be used in a
CC   disease-specific immunogen to protect a mammal against disease in
CC   particular melanomas. The peptides may also be used to screen a sample
CC   for the presence of an antigen with the same epitope, or with a different
CC   cross-reactive epitope.
XX
SQ   Sequence   9 AA;

  Query Match             84.4%;  Score 38;  DB 19;  Length 9;
  Best Local Similarity   77.8%;  Pred. No. 9.3e+05;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADINL 9
              |||||::||
Db          1 YLSGANLNL 9


RESULT 11
AAW70045
ID   AAW70045 standard; peptide; 9 AA.
XX
AC   AAW70045;
XX
DT   22-OCT-1998  (first entry)
XX
DE   CEA derived HLA-A2.1 binding peptide 2 (residues 605-613).
XX
KW   Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW   human leukocyte antigen; HLA; tumour associated antigen; cancer;
KW   antigen presenting cell; APC; immunogenic peptide; immune disorder;
KW   viral infection; AIDS; hepatitis; bacterial infection; malaria; CEA;
KW   fungal infection; tuberculosis; melanoma; carcinoembryonic antigen.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9833888-A1.
XX
PD   06-AUG-1998.
XX
PF   30-JAN-1998;   98WO-US01959.
XX
PR   31-JAN-1997;   97US-0036696.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Celis E,  Sette A,  Sidney J,  Southwood S,  Tsai V;
XX
DR   WPI; 1998-437445/37.
XX
PT   Production of antigen-specific cytotoxic T cells - by incubating
PT   immunogenic peptide(s) from antigen that binds class I major
PT   histocompatibility complex molecules with pre-treated antigen
PT   presenting cells
XX
PS   Example 6; Page 75; 104pp; English.
XX
CC   Sequences shown in AAW70044 to AAW70052 represent peptides derived from
CC   carcinoembryonic antigen (CEA). The peptides can bind to a human
```

```
CC   leukocyte antigen (HLA), HLA-A2.1 and are used to exemplify the method
CC   of invention of producing antigen-specific cytotoxic T cells (CTLs) in
CC   vitro. The method comprises contacting immunogenic peptides from an
CC   antigen that binds class I major histocompatibility complex (MHC)
CC   molecules with antigen presenting cells (APCs) pretreated with
CC   pretreatment growth factors, and incubating the APCs with purified CD8
CC   cells in the presence of at least 2 incubation growth factors, thereby
CC   producing antigen-specific CTLs. A method for specifically killing
CC   target cells in a human patient is also provided which comprises
CC   obtaining a fluid sample containing CTLs from a patient, contacting the
CC   cytotoxic T cells with APCs pretreated with pre-treatment growth
CC   factors, where the APCs comprise class I MHC molecules. The pretreated
CC   APCs are incubated with the cytotoxic growth factors, thereby producing
CC   activated CTLs which are contacted with a carrier to form a composition.
CC   The composition can then be administered to the patient. The activated
CC   CTLs can be used for treating cancers, immune disorders, viral
CC   infections, AIDS, hepatitis, bacterial infection, fungal infection,
CC   malaria or tuberculosis.
XX
SQ   Sequence   9 AA;

  Query Match             84.4%;  Score 38;  DB 19;  Length 9;
  Best Local Similarity   77.8%;  Pred. No. 9.3e+05;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADINL 9
              |||||::||
Db          1 YLSGANLNL 9


RESULT 12
AAY47655
ID   AAY47655 standard; Peptide; 9 AA.
XX
AC   AAY47655;
XX
DT   01-DEC-1999  (first entry)
XX
DE   Immunogenic peptide having a human leukocyte antigen binding motif #2266.
XX
KW   Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW   immune response; T cell activation; major histocompatibility complex;
KW   cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW   prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW   vaccine; immunisation.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9945954-A1.
XX
PD   16-SEP-1999.
XX
PF   13-MAR-1998;   98WO-US05039.
XX
PR   13-MAR-1998;   98WO-US05039.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Kubo RT,  Sidney J,  Celis E,  Grey HM,  Southwood S;
XX
DR   WPI; 1999-551214/46.
XX
PT   New immunogenic peptides with HLA binding motif, useful in treatment
PT   and diagnosis of cancers and viral diseases  -
XX
PS   Claim 1; Page 118; 150pp; English.
XX
CC   AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
CC   having a human major histocompatibility complex (MHC) Class I (also
CC   known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC   peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC   HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
```

```
CC   response against the antigen from which the peptide is derived.
CC   Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC   normally induced by an antigen in the form of a peptide fragment bound
CC   to a HLA molecule, rather than the intact foreign antigen itself, and
CC   are particularly important in tumour rejection and in fighting viral
CC   infections. The peptides are therefore useful therapeutically to treat
CC   or prevent viral infections and cancers in mammals (especially humans)
CC   e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC   They can be administered as vaccines to elicit an immune response in
CC   individuals susceptible or otherwise at risk of viral infection or
CC   cancer, or used to treat chronic or acute conditions. They are also
CC   useful diagnostically, and can be used to induce a cytotoxic T cell
CC   response, by contacting a cytotoxic T cell with the peptide e.g. to
CC   produce CTLs ex vivo for infusion back into a patient. The
CC   polynucleotides encoding the immunogenic peptides are also useful
CC   therapeutically and for immunisation as above.
XX
SQ   Sequence   9 AA;

  Query Match             84.4%;  Score 38;  DB 20;  Length 9;
  Best Local Similarity   77.8%;  Pred. No. 9.3e+05;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADINL 9
              |||||::||
Db          1 YLSGANLNL 9

RESULT 13
AAY09525
ID   AAY09525 standard; peptide; 9 AA.
XX
AC   AAY09525;
XX
DT   20-JUL-1999  (first entry)
XX
DE   Carcinoembryonic antigen peptide agonist CAP-1.
XX
KW   Carcinoembryonic antigen; CEA; human; agonist; antagonist;
KW   immune response; carcinoma; gastrointestinal; breast; pancreatic;
KW   bladder; ovarian; lung; prostatic; T cell proliferation; cancer;
KW   adoptive transfer therapy; autoimmune reaction; immunotherapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9919478-A1.
XX
PD   22-APR-1999.
XX
PF   22-SEP-1998;   98WO-US19794.
XX
PR   10-OCT-1997;   97US-0061589.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Barzaga E,  Schlom J,  Zaremba S;
XX
DR   WPI; 1999-326544/27.
XX
PT   Peptide agonists and antagonists of carcinoembryonal antigen
XX
PS   Claim 1; Page 53; 72pp; English.
XX
CC   The present invention describes peptides (A) that comprise agonists (Ia)
CC   or antagonists (Ib) of human carcinoembryonal antigen (CEA). (Ia) are
CC   used in vaccines to kill or inhibit carcinoma cells that express CEA or
CC   its epitopes, particularly for treating gastrointestinal, breast,
CC   pancreatic, bladder, ovarian, lung or prostatic carcinoma. They can also
CC   be used to proliferate T cells, e.g. from vaccinated subjects, for use
CC   in adoptive transfer therapy. (Ib) are used to inhibit CEA-specific
CC   immune responses, e.g. in vaccinated subjects, to prevent an autoimmune
CC   reaction to cancer immunotherapy (i.e. to prevent attack on normal but
```

```
CC   CEA-expressing cells). (Ia) are more active than native sequence (I) and
CC   generate a highly specific and systemic anti-CEA response. Cytotoxic T
CC   cells generated recognize both (Ia) and native CEA epitopes. The present
CC   sequence represents a specifically claimed example of (Ia).
XX
SQ   Sequence   9 AA;

  Query Match             84.4%;  Score 38;  DB 20;  Length 9;
  Best Local Similarity   77.8%;  Pred. No. 9.3e+05;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADINL 9
              |||||::||
Db          1 YLSGANLNL 9

RESULT 14
AAB13749
ID   AAB13749 standard; peptide; 9 AA.
XX
AC   AAB13749;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Peptide fragment # 1 from human CEA.
XX
KW   Human; T-cell; immune response; antigen; epitope; B7 family molecule;
KW   Leukocyte function-associated antigen-3; LFA-3;
KW   Intercellular adhesion molecule-1; ICAM-1; vaccine; immunotherapy;
KW   colon polyp; Crohn's disease; ulcerative colitis; breast lesion;
KW   tumour; CEA.
XX
OS   Homo sapiens.
XX
PN   WO200034494-A1.
XX
PD   15-JUN-2000.
XX
PF   12-NOV-1999;   99WO-US26866.
XX
PR   09-DEC-1998;   98US-0111582.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (THER-) THERION BIOLOGICS CORP.
XX
PI   Schlom J,  Hodge J,  Panicali D;
XX
DR   WPI; 2000-431307/37.
XX
PT   Novel recombinant vector useful as immunogens and vaccines for
PT   stimulating and enhancing immunological responses to target cells and
PT   antigens expresses multiple co-stimulatory molecules such as B7-1,
PT   LFA-3, ICAM-1  -
XX
PS   Claim 18; Page 35; 188pp; English.
XX
CC   Costimulatory molecules have important roles in T-cell activation and
CC   therefore the immune response. The present invention relates to
CC   recombinant vectors which comprise of foreign nucleic acid sequences
CC   encoding at least three costimulatory molecules: a B7 family molecule,
CC   Leukocyte function-associated antigen-3 (LFA-3, human CD58) and
CC   Intercellular adhesion molecule-1 (ICAM-1, CD54) and optionally a foreign
CC   gene encoding a target antigen or immunological epitope. The present
CC   sequence is one such target antigen used in the present invention. The
CC   present sequence is a tumour-associated antigen. The vector of the
CC   present invention would be useful for providing an enhanced immune
CC   response to the present target antigen. The vector of the present
CC   invention may therefore be useful in immunotherapy for treating or
CC   preventing diseases caused by viruses, bacteria, protozoans, parasites,
CC   premalignant cells and tumour cells. The recombinant vector can be used
CC   to treat or prevent preneoplastic or hyperplastic states such as colon
CC   polyps, Crohn's disease, ulcerative colitis and breast lesions.
XX
```

```
SQ   Sequence   9 AA;

  Query Match             84.4%;  Score 38;  DB 21;  Length 9;
  Best Local Similarity   77.8%;  Pred. No. 9.3e+05;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADINL 9
              |||||::||
Db          1 YLSGANLNL 9


RESULT 15
AAB82776
ID   AAB82776 standard; Protein; 9 AA.
XX
AC   AAB82776;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Carcinoembryonic antigen peptide.
XX
KW   Telomerase reverse transcriptase; hTRT; human;
KW   cytotoxic T lymphocyte; major histocompatibility complex; cancer;
KW   tumour; human leucocyte antigen; HLA-A2.1; vaccine;
KW   carcinoembryonic antigen.
XX
OS   Homo sapiens.
XX
PN   WO200160391-A1.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05143.
XX
PR   15-FEB-2000; 2000US-0182685.
PR   15-FEB-2001; 2001US-0182685.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Zanetti M;
XX
DR   WPI; 2001-536552/59.
XX
PT   Vaccine for initiating and enhancing a cytotoxic T lymphocyte response,
PT   for treating cancers or tumours or for inducing immune response against
PT   tumours, comprises a telomerase reverse transcriptase peptide    -
XX
PS   Example 1; Page 12; 52pp; English.
XX
CC   The present sequence is that of a carcinoembryonic antigen peptide
CC   comprising amino acid residues 571-579.  The peptide was used as a
CC   reference peptide in comparison with human telomerase reverse
CC   transcriptase (hTRT) HLA-A2.1+ restricted peptide p540 (see
CC   AAB82772) in a HLA-A2.1 binding/stabilisation assay.  The induction
CC   of CTL responses in vitro and in vivo, and the susceptibility to
CC   lysis of tumour cells of various origins by hTRT CTL suggest that
CC   hTRT could serve as a universal cancer vaccine for humans.  A
CC   claimed universal vaccine for treating tumours of any origin
CC   comprises at least 1 hTRT peptide.  The peptide is 7-15 amino
CC   acid residues in length and may be modified to enhance binding to
CC   the major histocompatibility complex.  Also claimed is a method for
CC   inducing and enhancing a CTL response against cancer cells, involving
CC   harvesting blood leucocytes, pulsing with hTRT, and contacting
CC   cancer cells with the pulsed leucocytes.  A method for targeting
CC   CTL to tumour cells is also claimed, and involves administering a
CC   hTRT peptide to a mammal, especially a cancer patient.
XX
SQ   Sequence   9 AA;

  Query Match             84.4%;  Score 38;  DB 22;  Length 9;
  Best Local Similarity   77.8%;  Pred. No. 9.3e+05;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLSGADINL 9
              |||||::||
Db          1 YLSGANLNL 9

Search completed: January 12, 2004, 14:28:23
Job time : 32.25 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         January 12, 2004, 14:24:44 ; Search time 25.75 Seconds
                                           (without alignments)
                                           90.193 Million cell updates/sec

Title:          US-09-529-121A-5
Perfect score:  48
Sequence:       1 YLSGACLNL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters:      775

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_23:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_virus:*
                13:  sp_vertebrate:*
                14:  sp_unclassified:*
                15:  sp_rvirus:*
                16:  sp_bacteriap:*
                17:  sp_archeap:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                    SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                      Description
 ------------------------------------------------------------------------
     1     19    39.6      9  4  Q15999                  Q15999 homo sapien
     2     17    35.4      7 12   Q66113                  Q66113 cherry leaf
     3     17    35.4      8 13   Q9PS69                  Q9ps69 gallus gall
     4     15    31.2      8  3  P87225                  P87225 saccharomyc
     5     15    31.2      8  6  Q9TRY3                  Q9try3 sus sp. ins
     6     14    29.2      7 13   O42564                  O42564 fugu rubrip
     7     14    29.2      8  2  Q9X3K1                  Q9x3k1 prochloroco
     8     14    29.2      8 13   P82079                  P82079 limnodynast
     9     14    29.2      9  6  Q9TRS0                  Q9trs0 oryctolagus
    10     14    29.2      9 10   Q9FSZ2                  Q9fsz2 cicer ariet
    11     14    29.2      9 11   O35953                  O35953 mus musculu
    12     13    27.1      7  2  Q8KMS3                  Q8kms3 klebsiella
    13     13    27.1      8  4  Q9Y4X6                  Q9y4x6 homo sapien
    14     13    27.1      8  6  O02831                  O02831 oryctolagus
    15     13    27.1      9  2  Q47063                  Q47063 escherichia
    16     13    27.1      9  4  Q99887                  Q99887 homo sapien
```

```
    17     13    27.1      9 11  Q9QZA8                  Q9qza8 mus musculu
    18     13    27.1      9 12  Q69473                  Q69473 human herpe
    19     12    25.0      8  2  P77556                  P77556 escherichia
    20     12    25.0      8  4  Q15901                  Q15901 homo sapien
    21     12    25.0      8  4  Q9UMH9                  Q9umh9 homo sapien
    22     12    25.0      8  4  Q8IUB8                  Q8iub8 homo sapien
    23     12    25.0      8  7  Q95213                  Q95213 oryctolagus
    24     12    25.0      8 13   Q91098                  Q91098 manorina me
    25     12    25.0      8 13   Q90498                  Q90498 erythrura g
    26     12    25.0      8 13   P82082                  P82082 limnodynast
    27     12    25.0      8 13   P82083                  P82083 limnodynast
    28     12    25.0      9  2  Q9R635                  Q9r635 chlamydia t
    29     12    25.0      9  4  Q9UCN5                  Q9ucn5 homo sapien
    30     12    25.0      9  5  Q27396                  Q27396 babesia bov
    31     12    25.0      9  6  Q28112                  Q28112 bos taurus
    32     12    25.0      9  6  Q9TRW2                  Q9trw2 oryctolagus
    33     12    25.0      9  8  Q94NB1                  Q94nb1 microcebus
    34     12    25.0      9  8  Q94NB2                  Q94nb2 microcebus
    35     12    25.0      9  8  Q94NA9                  Q94na9 daubentonia
    36     12    25.0      9  8  Q94NB0                  Q94nb0 microcebus
    37     12    25.0      9  8  Q31653                  Q31653 anser caeru
    38     12    25.0      9 10   Q9FEC0                  Q9fec0 hordeum vul
    39     12    25.0      9 12   Q9E1U7                  Q9e1u7 hepatitis b
    40     12    25.0      9 15   Q8UTD7                  Q8utd7 human immun
    41     11    22.9      7  2  Q8GL12                  Q8gl12 borrelia bu
    42     11    22.9      7 11   Q8K3H6                  Q8k3h6 rattus norv
    43     11    22.9      7 12   Q67113                  Q67113 influenzavi
    44     11    22.9      7 12   Q9YQ10                  Q9yq10 transmissib
    45     11    22.9      8  2  Q9RQ49                  Q9rq49 buchnera ap

                                ALIGNMENTS


RESULT 1
Q15999
ID   Q15999      PRELIMINARY;      PRT;     9 AA.
AC   Q15999;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   C-Kit protooncogene (Fragment).
GN   KIT.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92020918; PubMed=1717985;
RA   Giebel L.B., Spritz R.A.;
RT   "Mutation of the KIT (mast/stem cell growth factor receptor)
RT   protooncogene in human piebaldism.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:8696-8699(1991).
DR   EMBL; S58152; AAB19972.1; -.
DR   EMBL; S58145; AAB19972.1; JOINED.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  875 MW;  D32C74087041AEBD CRC64;

  Query Match             39.6%;  Score 19;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3; Conservative    0; Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GAC 6
              |||
Db          1 GAC 3


RESULT 2
Q66113
ID   Q66113      PRELIMINARY;      PRT;     7 AA.
AC   Q66113;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
```

```
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   C-terminus of the viral replicase (Fragment).
OS   Cherry leaf roll virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Comoviridae;
OC   Nepovirus.
OX   NCBI_TaxID=12615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Walnut;
RA   Borja M.;
RL   Thesis (1992), Biologia Molecular y Virologia Vegetal, CIT-INIA.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Walnut;
RX   MEDLINE=96124520; PubMed=8560786;
RA   Borja M., Sanchez F., Rowhani A., Bruening G., Ponz F.;
RT   "Long, nearly identical untranslated sequences at the 3' terminal
RT   regions of the genomic RNAs of cherry leafroll virus (walnut
RT   strain).";
RL   Virus Genes 10:245-252(1995).
DR   EMBL; Z34265; CAA84019.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  800 MW;  7417672EBDC6D740 CRC64;

  Query Match            35.4%;  Score 17;  DB 12;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ACL 7
              |||
Db          3 ACL 5


RESULT 3
Q9PS69
ID   Q9PS69      PRELIMINARY;      PRT;     8 AA.
AC   Q9PS69;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   LOW density lipoprotein receptor-related protein (Fragment).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92011685; PubMed=1918027;
RA   Stifani S., Barber D.L., Aebersold R., Steyrer E., Shen X., Nimpf J.,
RA   Schneider W.J.;
RT   "The laying hen expresses two different low density lipoprotein
RT   receptor-related proteins.";
RL   J. Biol. Chem. 266:19079-19087(1991).
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  846 MW;  C007272DD865BAAA CRC64;

  Query Match            35.4%;  Score 17;  DB 13;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 8.3e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGACL 7
              ||| |
Db          3 SGALL 7


RESULT 4
P87225
ID   P87225      PRELIMINARY;      PRT;     8 AA.
AC   P87225;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
```

```
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   GIN11 protein (Fragment).
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wedler H., Wedler E., Scharfe M., Wambutt R.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   MIPS;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z73169; CAA97518.2; -.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  1019 MW;  4E21A9C449D5B73B CRC64;

  Query Match            31.2%;  Score 15;  DB 3;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YLS 3
              |||
Db          1 YLS 3


RESULT 5
Q9TRY3
ID   Q9TRY3      PRELIMINARY;      PRT;     8 AA.
AC   Q9TRY3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Insulin-like growth factor-binding protein-6, IGFBP-6 (Fragment).
OS   Sus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9826;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92049376; PubMed=1719383;
RA   Shimasaki S., Gao L., Shimonaka M., Ling N.;
RT   "Isolation and molecular cloning of insulin-like growth factor-binding
RT   protein-6.";
RL   Mol. Endocrinol. 5:938-948(1991).
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  850 MW;  9FB2CEA37EA7687D CRC64;

  Query Match            31.2%;  Score 15;  DB 6;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 SGAC 6
               :| |
Db          1 AGPC 4


RESULT 6
O42564
ID   O42564      PRELIMINARY;      PRT;     7 AA.
AC   O42564;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   Truncated voltage-gated sodium channel alpha subunit (Fragment).
GN   SCN8A.
OS   Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
```

```
OC   Tetradontoidea; Tetraodontidae; Takifugu.
OX   NCBI_TaxID=31033;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97442476; PubMed=9295353;
RA   Plummer N.W., McBurney M.W., Meisler M.H.;
RT   "Alternative splicing of the sodium channel SCN8A predicts a truncated
RT   two-domain protein in fetal brain and non-neuronal cells.";
RL   J. Biol. Chem. 272:24008-24015(1997).
DR   EMBL; U97673; AAB80916.1; -.
KW   Ionic channel.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  730 MW;  75B72EA2C73772A0 CRC64;

  Query Match             29.2%;  Score 14;  DB 13;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 CLN 8
              ||:
Db          5 CLS 7


RESULT 7
Q9X3K1
ID   Q9X3K1      PRELIMINARY;      PRT;     8 AA.
AC   Q9X3K1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   Cytochrome b (Fragment).
GN   PETB.
OS   Prochlorococcus sp.
OC   Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC   Prochlorococcus.
OX   NCBI_TaxID=1220;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Urbach E., Chisholm S.W.;
RT   "Genetic diversity in Prochlorococcus populations flow cytometrically
RT   sorted from the Sargasso Sea and Gulf Stream.";
RL   Limnol. Oceanog. 43:1615-1630(1998).
DR   EMBL; AF070193; AAD23233.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  799 MW;  10376865B72866D3 CRC64;

  Query Match             29.2%;  Score 14;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |||
Db          4 LSG 6


RESULT 8
P82079
ID   P82079      PRELIMINARY;      PRT;     8 AA.
AC   P82079;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DYNASTIN 1.
OS   Limnodynastes interioris (Giant banjo frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Limnodynastinae; Limnodynastes.
OX   NCBI_TaxID=30362;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=TIBIAL GLAND;
RA   Raftery M.J., Bradford A.M., Bowie J.H., Wallace J.C., Tyler M.J.;
RT   "Peptides from Australian frogs. The structure of the dynastins from
RT   the banjo frogs Limnodynastes interioras, Limnodynastes dumerilii and
RT   Limnodynastes terraereginae.";
RL   Aust. J. Chem. 46:833-842(1993).
CC   -!- MASS SPECTROMETRY: MW=729; METHOD=FAB.
KW   Amphibian skin.
SQ   SEQUENCE   8 AA;  729 MW;  7C28772865B72728 CRC64;

  Query Match             29.2%;  Score 14;  DB 13;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |||
Db          3 LSG 5


RESULT 9
Q9TRS0
ID   Q9TRS0      PRELIMINARY;      PRT;     9 AA.
AC   Q9TRS0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Calcyclin-associated protein, CAP50=CA2+/phospholipid-binding protein
DE   L-7 fragment. (Fragment).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92250478; PubMed=1533622;
RA   Tokumitsu H., Mizutani A., Minami H., Kobayashi R., Hidaka H.;
RT   "A calcyclin-associated protein is a newly identified member of the
RT   Ca2+/phospholipid-binding proteins, annexin family.";
RL   J. Biol. Chem. 267:8919-8924(1992).
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1010 MW;  64E419C44865B72B CRC64;

  Query Match             29.2%;  Score 14;  DB 6;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |||
Db          3 LSG 5


RESULT 10
Q9FSZ2
ID   Q9FSZ2      PRELIMINARY;      PRT;     9 AA.
AC   Q9FSZ2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   Hypothetical 1.0 kDa protein (Fragment).
OS   Cicer arietinum (Chickpea) (Garbanzo).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.
OX   NCBI_TaxID=3827;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Castellana; TISSUE=Etiolated epicotyl;
RA   Dopico B., Jimenez T., Labrador E.;
RT   "cDNA clones expressed in etiolated Cicer arietinum epicotyls.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ299069; CAC10216.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  990 MW;  9441BDDAA7272EBE CRC64;
```

```
  Query Match             29.2%;  Score 14;  DB 10;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 LSGAC 6
              |  ||
Db          4 LLDAC 8


RESULT 11
O35953
ID   O35953      PRELIMINARY;      PRT;     9 AA.
AC   O35953;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Truncated voltage-gated sodium channel alpha subunit (Fragment).
GN   SCN8A.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RIII;
RX   MEDLINE=97442476; PubMed=9295353;
RA   Plummer N.W., McBurney M.W., Meisler M.H.;
RT   "Alternative splicing of the sodium channel SCN8A predicts a truncated
RT   two-domain protein in fetal brain and non-neuronal cells.";
RL   J. Biol. Chem. 272:24008-24015(1997).
DR   EMBL; U97672; AAB80914.1; -.
DR   MGD; MGI:103169; Scn8a.
KW   Ionic channel.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  898 MW;  22D92865B735B737 CRC64;

  Query Match             29.2%;  Score 14;  DB 11;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSG 4
              |||
Db          5 LSG 7


RESULT 12
Q8KMS3
ID   Q8KMS3      PRELIMINARY;      PRT;     7 AA.
AC   Q8KMS3;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Putative MerR2 protein.
GN   MERR2.
OS   Klebsiella sp. LS13-39.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Klebsiella.
OX   NCBI_TaxID=143776;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LS13-39;
RX   MEDLINE=21604134; PubMed=11763242;
RA   Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,
RA   Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,
RA   Yurieva O.V., Nikiforov V.G.;
RT   "Mercury resistance transposons of Gram-negative environmental
RT   bacteria and their classification.";
RL   Res. Microbiol. 152:811-822(2001).
DR   EMBL; AJ302776; CAC82975.1; -.
SQ   SEQUENCE   7 AA;  608 MW;  6DC1B5BDD87DD6F0 CRC64;

  Query Match             27.1%;  Score 13;  DB 2;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LSGA 5
              ::||
Db          1 MAGA 4


RESULT 13
Q9Y4X6
ID   Q9Y4X6      PRELIMINARY;      PRT;     8 AA.
AC   Q9Y4X6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Nuclear LIM interactor (Fragment).
GN   NLI.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20108806; PubMed=10640831;
RA   Drechsler M., Schumacher V., Friedrich S., Wildhardt G., Giesler S.,
RA   Schroth A., Bodem J., Royer-Pokora B.;
RT   "Genomic structure, alternative transcripts and chromosome location of
RT   the human LIM domain binding protein gene LDB1.";
RL   Cytogenet. Cell Genet. 87:119-124(1999).
DR   EMBL; AJ243097; CAB45408.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  767 MW;  EE6EBDDEB862D5B6 CRC64;

  Query Match             27.1%;  Score 13;  DB 4;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 AC 6
              ||
Db          6 AC 7


RESULT 14
O02831
ID   O02831      PRELIMINARY;      PRT;     8 AA.
AC   O02831;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Pro alpha 1 type III collagen protein (Fragment).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96377339; PubMed=8783186;
RA   Metsaranta M., Kujala U.M., Pelliniemi L., Osterman H., Aho H.,
RA   Vuorio E.;
RT   "Evidence for insufficient chondrocytic differentiation during repair
RT   of full-thickness defects of articular cartilage.";
RL   Matrix Biol. 15:39-47(1996).
DR   EMBL; S83371; AAD14433.1; -.
KW   Collagen.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  1028 MW;  B859C7272EA77371 CRC64;

  Query Match             27.1%;  Score 13;  DB 6;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 CL 7
              ||
Db          4 CL 5
```

```
RESULT 15
Q47063
ID   Q47063      PRELIMINARY;      PRT;     9 AA.
AC   Q47063;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   URF 1.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=W;
RX   MEDLINE=85215599; PubMed=2987841;
RA   Takagi J.S., Ida N., Tokushige M., Sakamoto H., Shimura Y.;
RT   "Cloning and nucleotide sequence of the aspartase gene of Escherichia
RT   coli W.";
RL   Nucleic Acids Res. 13:2063-2074(1985).
DR   EMBL; X02307; CAA26175.1; -.
SQ   SEQUENCE   9 AA;  1061 MW;  9DE21EA5B9C72EA1 CRC64;

  Query Match            27.1%;  Score 13;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 CL 7
              ||
Db          3 CL 4

Search completed: January 12, 2004, 14:31:02
Job time : 26.75 secs
```